Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AI-Based Deepfake Detection for Fraud Prevention and Cybersecurity in U.S. Digital Banking Systems

MD Hasibur Rahman Murad[1], Kabir Hossain Tutul[2], Md Atiqur Rahman[3], Rakibul Hassan Bhuiyan[4], Saidul Islam[5], Farazi Ahsanul Islam[6], Md Mishal Mahmood[7]*, Shahed Ahmed Chowdhury[8]

[1]*School of Business Administration, Hunan University, China.*

[2]*College of Engineering, University of South Florida, Tampa, Florida,*

[3]*Master of Science in Information Technology, Washington University of Science and Technology, Virginia, USA.*

[4]*Bachelor of Science in Information Technology (BSIT), Washington University of Science and Technology,*

[5]*Master of Science in Information Technology, Washington University of Science and Technology, Virginia, USA.*

[6]*Master of Science in Information Technology (MSIT), Washington University of Science and Technology, Virginia,*

[*7]*Master of Science in Cybersecurity, Washington University of Science and Technology, Virginia, USA.*

[8]*M.S. in Information Systems Technologies, Wilmington University, New Castle, Delaware, USA.*

Corresponding Author: Md Mishal Mahmood

ARTICLE INFO

Received: 03 Nov 2024

Revised: 17 Dec 2024

Accepted: 28 Dec 2024

ABSTRACT

Purpose: The authors examine the impact and contributions of AI solutions to detect deepfakes on fraud prevention and cybersecurity measures in digital banking in the USA. Specifically it explored the relationship between AI-Driven Multimodal Deepfake Detection Capability (AMMDFC) on Digital Banking Fraud and Cybersecurity Resilience (DBFCR) and the relationship between AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) on DBFCR and between Explainable AI-Based Deepfake Detection Capability (XAIDDC) and DBFCR, with Digital Banking Fraud and Cybersecurity Resilience (DBFCR) serving as a mediator.

Design/Methodology/Approach: Quantitative and Cross sectional design were used. 510 bankers, cyber experts, fraud prevention, risk, IT and digital-finance professionals were sampled. The proposed models in terms of measurement and structure were analyzed in a PLS-SEM application with a bootstrapping process via the SmartPLS software program version 4.0.The proposed models regarding the measurement and structure were tested with the PLS-SEM with 1,000 bootstrapping, using the software program SmartPLS.

Findings: First the results indicated that AMMDFC, ARTDDC, and XAIDDC were found to have positive and significant effect on DTDE and DBFCR. DTDE also managed to make significant improvement for DBFCR. The model was successful in explaining 54.2% of means variation for DE of DT and 64.0% of means variation for DBFCR. Second, the three descriptions of AI were fully mediated through DTDE in terms of relation to DBFCR.Secondly, three descriptions of AI were linked to DBFCR all the way through DTDE.DTDE was partially responsible for all three AI skills and DBFCR. ARTDDC was more effective on DTDE highlighting the significance of timely detection in digital banking environment.

Theoretical Implications: The study statistically anchored the Resource-Based View, Dynamic Capabilities Theory, Organizational Information Processing Theory and Organizational Resilience Theory and also highlighted the translation of AI capabilities into cybersecurity, mediated by effective threat detection.

Practical Implications: Banks need to incorporate multimodal analysis, real-time detection capabilities, explainable AI features, human controls, and incident response protocols into their technology to improve their defenses against synthetic identities, voice cloning, and fabricated digital content.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Originality/Value: The study combined the knowledge within an organizational structure that shows direct and indirect links between the implementation and development of complementary deepfake-detection and digital banking resistance.

Keywords: Deepfake detection; Artificial intelligence; Digital banking; Fraud prevention; Cybersecurity resilience; Multimodal AI; Explainable AI; Real-time detection; Synthetic identity; PLS-SEM.

INTRODUCTION:

In the United States, digital identity and digital interactions are more vital to banks than ever before—and have fueled a new generation of advanced deepfake financial services frauds and cyberattacks. As generative AI advances, faces, videos, voice and identities may continue to become increasingly realistic, making it difficult for identity verification, authentication and fraud-detection systems to keep up. In the future, as generative AI becomes increasingly capable of creating real-looking faces, videos, voices, and identities, traditional methods of identity verification, authentication and fraud detection may struggle to keep pace. These capabilities can also be harnessed for a range of attacks in digital banking, such as: impersonating customers, creating synthetic identities and exploiting social engineering, defeating authentication, accessing fraudulent accounts and manipulating transactions in the remote transaction authorization process. Thus, recent research has focused on the use of deep learning artificial Intelligence (AI) for analyzing manipulation artifacts and inconsistencies of the visual, auditory, biometric, contextual and other forms of digital signals as a new capability of cyber security (Staynings, 2023). Further, the multimodal approach is used as deepfakes are known for their ability to impact multiple components of information, and real-time detection is becoming crucial to intercept the fraudulent communications, before the financial transaction or unauthorized entry are made. In the interim, explainable AI has become a matter of greater importance as algorithms become completely un-explainable and only high-profile financial institutions can demand detection results that can be understood, investigated, validated and activated to take possible measures for fraud/cybersecurity responses. However, recent advances in making ever more realistic and flexible synthetic media have kept enhancing the Going-Great-Times with the various synthesis methods and detection techniques for the generation of deepfake. It is relevant to discuss the possibility to evaluate the capability of AI for Multimodal Deepfake Detection, Real-Time Deepfake Detection and Explainable AI Deepfake Detection in this context so as to determine how the U.S. digital banks will become more resilient in the context of rapidly changing AI-based financial threats and risks.

With generative AI, fraudsters can create realistic AI generated faces, video manipulation, voice-generated text with AI, synthetic identities, and fabricated faces, voices, and identities that could be used to fool the digital bank's transaction system and facilitate financial gain cybercrimes, such as customer ID theft, social engineering, and synthetic identity fraud. Although there is immense adoption of AI-based fraud security and detection tools by the banks, the studies in this field indicate that deepfake detection is still very challenging because of deepfake generation methods' explosion, multimodal attacks, the requirement of real-time technology protection, generation systems' robustness, and the difficulty of seeing through an AI-based system because of the opacity of decisions made via AI. A capability for AI-Driven Multimodal Deepfake Detection (AI-DMDD), that integrates multiple modalities' datasets to detect inconsistencies in such synthetic interactions, may help to solve these problems, and AI-Enabled Real-Time Deepfake Detection Capability (AI-ERDDC) may help banks identify these deception attempts in time to prevent damage occurring before it does. Besides the need for Explainable AI-Based Deepfake Detection Capability (XAIDDC) in financial high stakes environment is huge as when interpretable results that could be validated, it could be utilized for investigative purposes and then an appropriate response could be taken for the threats identified by the algorithm; particularly in the financial high stakes environment where the detection takes place in such a high level of uncertainty, that the level of trust and credibility in the processes could be limited. The research on the three organizational AI features which have been mentioned above, however, has been largely isolated and each feature has been studied in isolation separately focusing on either the contributions of each feature to Digital Banking Fraud and Cybersecurity Resilience (DBFCR), or the individual capabilities of deepfake algorithms and explainability, and with comparatively little empirical research on how they can work together to achieve benefit and, if they do, work through Deepfake Threat Detection Effectiveness (DTDE). It is a very important problem because detection of voice

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

cloning/manipulated identities/synthetic media/ deepfakes will be possible thanks to technology and frauds can finally be detected before they become a threat to banking services (LOZONSCHI & BAKHAYA, 2023). So, the central question of this research is how a direct impact of AMDDDC and ARTDDC on DBFCR is explained, how DTDE explain the relationship between the DDDs and the DBFCR, and how this impact upon DBFCR is indirect via DTDE within the U.S. digital banking system.

Although the technology to identify DeepFakes is growing quickly, there are critical gaps in theory, empirical evidence, and industry practice that have yet to be filled to effectively understand how DeepFakes can contribute to preventing fraud and cyber resiliency in today's U.S. digital banking systems. While numerous existing studies on deepfake detection have proposed methodologies that mostly focus on algorithmic accuracy and testing errors, improvements, or alternative algorithms, gaps in research and writing generally have overlooked the various ways in which deepfake detection can be operationalized as an organizational, cybersecurity-enabling activity, where technological assets are integrated into ways of achieving broader fraud prevention and resilience goals (Mendoza, 2024). An explanation or a theoretical understanding, therefore, only partly exists to connect the digital banking fraud and cybersecurity resilience (DBFCR), Deepfake Threat Detection Effectiveness (DTDE), AI-Driven Multimodal Deepfake Detection Capability (AMDDFC), and the AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) and the AI-Extended Explainable AI-Based Deepfake Detection Capability (XAIDDC). While separately, recent studies have highlighted the potential of multimodal fusion, real-time AI, and explainable AI to provide better deepfake threat detection, the research streams have been limited, and to the best of our knowledge, nothing simultaneously tests the three capabilities by showing how improving deepfake threat detection directly improves organizational fraud and cybersecurity resilience capabilities. Additionally, current research focuses mainly on proposed architectures, assessed datasets under lab conditions, either caused-by and either caused-by-qualities analysis of single audio or visual modality, or on technical performance aspects, rather than on causal and mediation relationships between variables at the organizational level, that is, particularly when applied in operational conditions such as banking environments (Schiliro, 2024). Remote onboarding, mobile transactions, digital-KYC, videoconferencing, and voice onboarding- and ID verification-aspects of U.S. digital banking are all mentioned as points of vulnerability for digital-banking services that should be closely monitored as key areas for the emergence of deepfake attacks and synthetic identity fraud, while recent financial-sector literature is increasingly pointing to the need for appropriate and scalable AI mechanisms that are real-time, explainable, and governance-aware and can transpose detection results into downstream effects, such as risk-response, as a potential area for risk (Sailaja, 2023). Based on the gaps identified in the theoretical and empirical literature, as well as the U.S. banking sector, this study incorporates all three of the above (AMDDFC, ARTDDC, and XAIDDC) in one framework and investigates how they directly affect the DBFCR, their impact on DTDE, the impact of DTDE on DBFCR, and how they mediate in the relationship between the depth of AI-based deepfake detection ability and DBFCR.

The study value lies in the opportunity it provides to raise awareness and encourage wider adoption of the power of AI to develop deepfake detection moving from merely a narrow technical task of determining media class and becoming a comprehensive organizational approach to a fraud prevention and cybersecurity resilient effort in digitally banked systems in the United States. The overall goal of the study is to gain understanding of the role of the proposed capabilities in Digital Banking Fraud and Cybersecurity Resilience (DBFCR) and more specifically how the proposed Deepfake Threat Detection Effectiveness (DTDE) is realized to attain better security outcomes. With the ability to harvest high-quality synthetic voices in order to generate new identities, create synthetic identities, manipulate audiovisual information and impersonate existing customers, financial institutions are in need of fresh solutions that can identify attack and react to them in financial services digital identity processes, such as remote onboarding, transaction authorization and communications with customers. Complementary information has been found to be used in conjunction with multimodal analytics to facilitate the identification of deepfakes used for fraud, and real-time AI could be used more effectively when it sought to identify adversarial behaviour before it was committed (Horowitz et al., 2022). Explainability is also associated with further cybersecurity advantages, as it enables identification of Cyber Threat Assessments (CTAs) that are easily understood and can be utilized by cyber analysts and decision-makers to enhance investigations, validation, accountability and appropriate Cyber Security responses in highly contaminated financial spaces. In the context of cybersecurity, the study aims to answer the following two questions: does better deepfake-detection ability make banks more effective at spotting impersonation

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

and synthetic-types of fraud; and will better DTDE lead to greater aspects of banks' fraud and cyber incident prevention, response and recovery? It is a purpose integrated and recent research has shown, with secure financial cybersecurity, there is far more than just detection required; namely, it requires real-time threat identification and adaptive response plus the need for interpretable AI. For this reason, the study aims to also provide theoretical and empirical insights concerning the ability of AI to identify all forms of deepfakes, as well as ABDD or ADDC as concrete examples of how the DBFS, AMDDFC, ARTDDC and XAIDDC are able to augment DTDE and, eventually, DBFCR, thus contributing to more reliable, adaptive and resistant cyber banking security architectures.

The novelty of this study lies beyond the scope of the aging algorithm and media-centric mindset of traditional suspects—the “old guard” in deep fakes—forensics—providing a contemporary lens for examining the role of AI in deepfake detection and accuracy based on the new organizational cyber-security/capability lens and the specific context of the study: the United States of America and digital banking systems. Yet, as reported in these related works, there are relatively few connections to these concepts in the literature and limited interaction with each other in the technical and non-technical aspects of using AI in their applications. For example, multimodal approaches are found to be helpful in fusing complementary signals and real-time approaches are highlighted to cope with adversarial threats which change rapidly, and explainability, which is differentiated from transparency in financial services, is recognized to be essential to ensure transparent and accountable AI decisions (Brooks, 2024). The authors, therefore, suggest that the AI-Driven Multimodal Deepfake Detection Capability (MMDDFC), AI-Enabled Real-Time Deepfake Detection Capability (RTDDFC), and Explainable AI-Based Deepfake Detection Capability (XAIDDC) are three complementary and distinct organizational capabilities examined and considered within the context of Digital Banking Fraud and Cybersecurity Resilience (DBFCR). Most notably, it has an explanatory mediate, called Deepfake Threat Detection Effectiveness (DTDE), which examines the impact of these AI technologies on the level of resilience as well as how increasing threat detection efficacy of deceptive identity & voice, cyber-bot, and manipulation of digital interactions would boost fraud-prevention and cybersecurity results. This mediation perspective aims to solve the issues of multimodal analysis of financial cybersecurity needed for financial tasks, not to be only one more addition of technologies, alongside with real-time adaptivity and detection, explainability and operational responses for threats (Anagha et al., 2023). The novelty is three sets of novel AI-based deepfake detection tools (DTDE, DBFCR) are being tested and subjected to examination within one mediated organizational compositions specifically in the context of digital banking in the U.S.

In digital banking, the power of remote banking, digital-KYC, mobile payments and digital-identity verification, has hit shoulder-to-shoulder grounds with easy availability of generative Artificial Intelligence (AI) that can generate synthetic identities, clone voices and images, and even create appropriate realistic ‘togetherness’ to build effective audiovisual cyborgs, providing functionality for the set-up of financial fraud and cyberattacks (Shrivastava et al., 2024). Encoding a single, conventional security method could be insufficient as synthetic content continues to evolve, presenting increased demands for the multimodal, real-time detection and interpretable AI features of contemporary financial cybersecurity measures. The deployment of accurate AI models is also not sufficient because it is critically important that models are deployment-ready, timely, and that the detection process is adaptive to the AI’s model, transparent and can also assist in making security decisions during the operational phase of a detector's deployment. AI-DMDP research is thus essential as it will identify various types of manipulation signals, elevate the chance of intercepting AI-driven deepfakes prior to their arrival at the consumption point, and will aid security personnel to comprehend and react to signals generated by AI. Additionally, to investigate the role of Deepfake Threat Detection Effectiveness (DTDE) as a mediator is key to understand the link between the investment in these technologies and improved results in Digital Banking Fraud and Cybersecurity Resilience (DBFCR), as there is currently no evidence of a relationship between the former and the latter. The key message of this investigation is particularly relevant in the context of digital banking in the U.S. where the demand with the advent of AI-fraud responses requires a strong unity and collaboration of detection, identity protection, cybersecurity, governance and operational resiliency (OR) (Saalbach, 2023). The study therefore aims to offer empirical evidence to better inform investment in AI security and bolster banks' ability to detect, deter, defend against, and recover from new threats enabled by deepfakes.

This research is important because it provides means to develop theory and practice that can enhance cybersecurity fraud prevention as well as cybersecurity resilience in U.S digital banking through investigating how complementary

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AI-based deepfake detection capabilities can do so. Financial cybersecurity has become an essential focus recently, with deep fake and synthetic identity attacks posing new challenges on financial institutions, and explainable, adaptive and resilient financial fraud detection through AI, recent research has shown. The study adds to the literature by incorporating and demonstrating the entire suite of AI-Driven Multimodal Deepfake Detection Capability (AMDDFC), AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) and Explainable AI-Based Deepfake Detection Capability (XAIDDC) as a suite within one framework and by articulating the Deepfake Threat Detection Effectiveness (DTDE) that connects these capabilities to Digital Banking Fraud and Cybersecurity Resilience (DBFCR). This paper extends the current research done on deepfake detection, explainability, fraud analytics, and cybersecurity resilience, which generally and often are considered separately in the literature. In the real world, the results might inform U.S. banks, digital banking companies, cybersecurity managers and fraud investigators, as well as technology security operations teams and risk managers and technology developers with clear guidance on which AI functions are most critical to detect and combat threats through deepfake technology. Finally, the evidence for explainable and operationally effective AI could also provide inputs to regulators and policymakers developing guidance on responsible governance of AI in the financial sector, such as cybersecurity. Last, the integrated capability–detection–resilience framework can serve as the basis for further empirical work and those who are using digital banking indirectly are likely more easily protected in their identity, are better protected from fraud, can use digital banking safely and are provided with more resilient digital banking services.

An interesting aspect of the study is exploring the relationships between the study variables since deepfake-hempstered banking fraud is considered a multi-dimensional banking cybersecurity challenge where aspects of breadth, speed of detection, interpretability of the detection, effectiveness of threat identification and organizational resilience are logically connected with detection breadth and interpretability in particular having rarely been studied together under one empirical framework. Illustrative examples show how multimodal AI systems can proactively combine complementary information such as audio, visual, biometric, and contextual to greatly improve the detection of advanced deepfakes, and how real-time AI can be utilised to provide financial institutions with a view into emerging fraudulent behavior as it occurs in digital interactions, as opposed to reacting after loss has occurred (Hämäläinen, 2024). Concurrently, with such detection decisions being easily interpretable and actionable by both financial-segment decision makers and cyber security professionals, explainable AI (XAI) is gaining importance (Reddy, 2023). The direct connections between the proposed newly defined DC capabilities – the AI-Driven Multimodal Deepfake Detection Capability (AMDDFC), the AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) and the Explainable AI-Based Deepfake Detection Capability (XAIDDC) to Digital Banking Fraud and Cybersecurity Resilience (DBFCR) and their relation to the newly proposed Digital Security Level (DSL) are worth studying to better understand these connections. Another important aspect of the relationship between DTDE and DBFCR is the understanding that the ability to identify and filter out advanced AIs does not necessarily mean that the organisation is resilient to them. If, for instance, the advanced AIs are not "identified" either and used for deepfakes or "deep fake" digital interaction, or an equivalent "fighting advanced AIs" security solution is not put in place, the whole situation is threatened. This novelty of these relationships thus extends beyond measures of the technical performance of detecting technologies, such as recognition accuracy, to "hard" measures, to DTDE directly linking three complementary capacities of AI and banking resilience more broadly. In total, the seven direct connections among these various circuits (between AMDDFC and DBFCR; ARTDDC and DBFCR; XAIDDC and DBFCR; between AMDDFC and DTDE; ARTDDC and DTDE; and XAIDDC and DTDE; and between DTDE and DBFCR), plus three intervening circuits (DTDE to DBFCR) provide a full picture of the potential for transferring AI-based deepfake detection criteria to enhance the fraud prevention and cybersecurity resiliency of U.S. digital banking systems.

Addressing the potential of increasingly sophisticated synthetic-media threats, this study aims to empirically evaluate the benefits of AI-based deepfake detection technologies to fraud prevention and cybersecurity resilience in U.S. digital banking systems, focusing on using multimodal, real-time and explainable method. The direct impact of the three capabilities (AI-Driven Multimodal Deepfake Detection Capability (AMDDFC), AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) and Explainable AI-Based Deepfake Detection Capability (XAIDDC) on Digital Banking Fraud and Cybersecurity Resilience (DBFCR) and the impact on Digital Banking Fraud and Cybersecurity Resilience (DBFCR) by the three capabilities on Deepfake Threat Detection Effectiveness (DTDE) and

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

the impact on Deepfake Threat Detection Effectiveness (DTDE) by the three capabilities on Digital Banking Fraud and Cybersecurity Resilience (DBFCR). Specifically, this study examines whether, in the tested model of DBFCR, only engages with the three dimensions of efficacy—expertise, UI, and trust—have been mediated by degree of engagement with DTDE; and whether a tested engagement with the three efficacy dimensions to DBFCR is mediated by the three cybersecurity capacities of AI: expertise, UI, and trust; to explicate how the three cybersecurity capacities of AI technology might be translated into overall dimensions of DBFCR at the organizational level. These goals are in line with the new research that suggests that deploying multiple "one-to-one" defense systems might not be the best approach, and instead, a few systems as one, with multimodal analysis, rapid threat identification, explainability and effective defenses at the operational level, could be better (Dewani et al., 2022).The broad objective of the study is to develop and test the integrated framework to elucidate how the complementary capabilities of AI supports alerts for deepfakes and frauds, and enhance the resilience, response and recovery capabilities of U.S. digital banks to emerging AI-based cyber-attacks.

THEORETICAL FOUNDATION:

This study has conceptual roots in Resource Based View (RBV) and Dynamic Capabilities Theory (DCT). According to RBV, if resources and capabilities in the organisation (or firms) are rare, non-substitutable and organized, they can provide better results. Here three specific technological and analytical capabilities exist to help protect U.S. banks' digital operations: AI-Driven Multimodal Deepfake Detection Capability (AMDDFC); AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC); and Explainable AI-Based Deepfake Detection Capability (XAIDDC). AMDDFC signifies the skill-set needed to integrate facial, video, voice, image, biometric and contextual signals into digital interactions; ARTDDC signifies skills needed for rapid detection of synthetic manipulation in digital interactions; and XAIDDC represents the skills needed to provide meaningful evidence for detection decisions. Multimodal analysis, speedy detection powered by AI and the explainability are crucial to countering deeply knotted deepfake and synthetic fraud, as backed by contemporary cyber research (Kolade et al., 2024). The constant monitoring of threats, the selection of the appropriate ones, and the shifting of the structure and procedures of the organization (Teece et al., 1997, Teece, 2007) is also an inherent part of an organization in a rapidly changing threat environment. Overall, complimenting each other, the three AI capabilities are like dynamic security attributes that can assist banks in dealing with the evolving landscape of deepfake fraud.

The DTDE can be theoretically explained that uncertainty and complexity of environments trigger organizational changes towards decision effectiveness by enhancing its information processing capacity. If AI can be used to create a deep-fake of a face, voice, body, or voice and face, this brings damaging information ambiguity: authenticity decision making might involve considering multiple sets of disdiverse and heterodox cues and evidence (visual and vocal, body and vocal, voice and face, visual and body, visual and face, and body and face) simultaneously. AMDDFC provides additional information for threat assessments, ARTDDC speeds up processing of the information and XAIDDC brings the detection information into a more readable, processable, and actionable format. DTDE is the type of information processing in between these technological abilities and what is effectively being achieved, which is the detection of synthetic identities, voice cloning, tampered audiovisual information and attempts at impersonation. These theoretical discussions are echoed in the literature, which emphasizes the importance of multimodal thinking for the real-time detection of increasingly complex frauds and cybersecurity attacks, as well as the importance of explainable decision support in these circumstances (Geradts & Franke, 2023).

The ultimate goal would be Digital Banking Fraud and Cybersecurity Resilience (DBFCR), which is consistent with Organizational Resilience Theory's perspective of an organization's ability to "anticipate threats, face disruptions, adjust responses and recover, while continuing to provide essential products and services" (Engvall, 2024). When applied to digital banking, DBFCR is more than just successful deep fake classification; it's a testament to the bank's ability to prevent fraud, defend against AI-driven attacks, react well, recovery from disruptions, and continue to maintain secure digital services. The current financial cybersecurity discussion also points to the use of Artificial Intelligence (AI) in risk detection and prevention of cyber frauds must not solely be a detection, but must also be an adaptive and resilient security system in an organization/firm. It then makes sense that a consistent theoretical pathway exists in the proposed model from the strategic capabilities and strategic actions of resilient AI systems (AMDDFC, ARTDDC), to DTDE bolstering the threat-information processing of organizational actions (XAIDDC),

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

to DBFCR. This is further supported by a theoretical basis which enables the analysis of the direct effects of the three AI features to DBFCR and indirect effects through DTDE.

Hypotheses Development:

H1: AI-Driven Multimodal Deepfake Detection Capability (AMDDFC) positively influences Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

AI Multimodal Deepfake Detection Capability helps banks to use multiple information modalities: facial, video, voice/audio, image, biometric, and contextual—rather than just one information modality. The ability to produce multi-modal attacks is especially pertinent, as it can be used to attack multiple modalities of the aforementioned in the same attack and thus outsmart the security based on individual signals. Recently, it has been revealed that multimodal methods can benefit deepfake detection or synthetic fraud by utilizing complementary information and capitalizing on the discrepancy in the information between different modalities which is only evident in the multimodal analysis. In the Resource Based View, these analytic resources of the bank are strategic resources which can be utilized in applying to digital operations protection with AI and heterogeneous security data.

Within the realm of resilience, a multimodal intelligence could enable a bank to resist the threat posed by deep-fake fraud, boost risk resilience and improve their understanding by providing additional evidence for decisioning actions. In recent times, integrated and multimodal defense configured in the context of combating synthetic media attacks is being considered. In this way, moneylenders having solid AMDDFC will have increased capability in battling identification manipulation and secure electronic banking activities in the context of threat from AI instruments. Thus, AMDDFC can be expected to positively impact DBFCR.

H2: AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) positively influences Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

The benefits of deep fakes to cybersecurity starts with being able to accurately detect a deepfaked interaction – if one can be identified before an unauthorized access or fraudulent transaction is undertaken. Discovering anomalous content in digital onboarding, Authentication, Transactions processing and Customer interactions is an area that banks can explore using ARTDDC to shorten the time delay between the anomalous content being perceived by the customer in the digital onboarding process, to the moment the bank's representative acts on it. Newer research in the financial sector focuses heavily on real-time fraud analytics, as well as real-time fraud detection enabled by AI technology, to improve threat response capabilities when dealing with new threats.

Dynamic Capabilities theory also suggests that the characteristics of dynamic environments require a responsiveness or skill that helps an organization detect the threat and respond very fast with an appropriate capability (Pavel, 2023). ARTDDC demonstrates this sensation, because when the deepfake attack occurs, it can be detected in real time / near real, leading to challenges to the authentication, controls of the transactions, or escalation / investigation of the attack. With earlier detection of exposure to deepfake fraud it is possible to reduce fraud exposure and secure digital banking continuity. Therefore, ARTDDC is expected to help to increase the DBFCR.

H3: Explainable AI-Based Deepfake Detection Capability (XAIDDC) positively influences Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

While advanced AI tools can identify fake content, artificially ambiguous predictions could reduce a cybersecurity professional's ability to understand the "how" and "why" behind an identified manipulation.While AI is powerful enough to generate notations of highly-skilled and crafted forms of manipulation, it's not always the best type to help the cybersecurity professional understand how and why an interaction has been flagged as manipulated. To overcome this advantage, XAIDDC offers interpretable evidence that is attached to detection findings and enables fraud analysts and cybersecurity teams to comprehend, verify and act on AI-generated alerts. In high-stakes financial contexts, like where security choices could impact transactions, identities, access to accounts, and regulatory duties, explainability is especially significant. Thus, explainability becomes an integral part of trustworthy financial AI and fraud-risk management, as it has recently been found to be one of the elements of financial trustworthiness (Nibigira et al., 2024).

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Additionally, from an information processing perspective, organizational information security performance relies on converting information into useful knowledge for organizational decision making (Nibigira et al., 2024). In order to make this transformation more interpretable and actionable for human decision makers, explainable detection enhances the use of algorithms. XAIDDC can then speed up investigations, help make better decision-making when it comes to fraud responses, achieve accountability, and enhance the coordination of automated with human security actors. These features can enhance the capabilities of banks to counteract and rebound from fraud situations triggered by AI, fostering a favorable XAIDDC-DBFCR dynamic.

H4: AI-Driven Multimodal Deepfake Detection Capability (AMDDFC) positively influences Deepfake Threat Detection Effectiveness (DTDE).

Deepfake attacks can be on multiple different information channels, such as facial appearance, movement of lips, speech, type of voice used, images, etc., and the contextual information. Multimodal capability enables a banking security to jointly analyze all these signals, which are heterogenous, and identify inconsistencies between modalities – independently of the fact that a single source might be compromised. Their recent work All-In-One Synthesis-COPs DeepFake Detection for Multimodal Deepfakes. Nnamdi et al. (2023) demonstrates that multimodal integration is crucial to enhance deepfake and synthetic-fraud detection, especially with the rise in sophisticated synthetic content creation abilities of generative technologies in multiple media formats.

OIPT proposes that if there is a high level of information uncertainty, then an organization will need it's information-processing capacity to be larger (Akpabio & Narad, 2024). AMDDFC does this by expanding the evidence available to consider in establishing the authenticity of an identity or digital interaction. Thus, cross-modal analysis can be applied to enhance face/face video recognition, voice cloning and synthetic identities, as well as impersonation of same using facial/video and vocal features. Theoretically better multimodal detection capabilities should support better DTDE as this should represent the effectiveness of the banks in detecting such threats.

H5: AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) positively influences Deepfake Threat Detection Effectiveness (DTDE).

However, for the sake of Deepfake Threat Detection Effectiveness, not only must banks detect the threats correctly, but fast enough. ARTDDC is able to offer a real-time or speedy analysis during potentially risky transactions with banking activity and so allows for identification of suspicious interactions which are synthetic in nature while they are happening. With the advent of real-time AI-powered attacks and their ability to evolve and outsmart traditional fraud defenses, real-time AI detection has emerged as a critical component. This ability to react in real time can thus minimize the chance of fraudulent interactions going 'under the radar' of the system, thereby preventing them from becoming recognized.

This also corresponds with Dynamic Capabilities Theory, with respect to organizational capability to identify threats as they arise in the environment. Rapid sensing is enabled by security systems to be able to sense abnormal audiovisual and/or identity signals during onboarding, authentication, transactions and customer communications in digital banking. Therefore, ARTDDC should aim towards augmenting the efficiency of the detection of deep fake identity and/or synthetic fraud and make it more timely, leading to better DTDE.

H6: Explainable AI-Based Deepfake Detection Capability (XAIDDC) positively influences Deepfake Threat Detection Effectiveness (DTDE).

Validating the threat and deciding which action to take requires more than simply detecting whether content being sent is real or fake, however — security personnel need to know what was behind the alarm to confirm that it's a threat and what to do in response. A deepfake classification report from XAIDDC includes usable data about the features, anomalies or inconsistencies that were responsible for the classification. From recent studies, it is emphasized that XAI has immense importance in Deepfake detection and financial fraud detection, as this is an area that introduces interpretability into the AI tools used to make security decisions, which may improve usability and trustworthiness of security decisions.

OIPT contributes to the quality of information for organizational interpretation and decision making by the way of explainability. If analysts are able to explain the reasoning behind the reported interaction, they can better confirm

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

the alert, eliminate false positives and provide their own Investigative or Preventative controls in response to the threat. Therefore, XAIDDC will make automated detection a more useful operation and will help the organisation further in successfully identifying threats, which are created with the help of a deepfake. Therefore, XAIDDC is hoped to have a positive impact on DTDE (Udeze et al., 2024).

H7: Deepfake Threat Detection Effectiveness (DTDE) positively influences Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

DTDE is the operation's efficiency of detecting these emerging frauds—synthetic identities, voice cloning, manipulated audiovisual evidence, and other forms of fraud involving AI—and deepfake impersonation.DTDE is the ability of the operation to detect these emerging forms of deepfake impersonation, synthetic identities, voice cloning, manipulated audiovisual evidence and other AI-based fraud. Effectively detecting it is a first line of defense as threats that have gone unnoticed are hard to contain or mitigate. In recent times, banking cybersecurity research has underscored the pivotal role of AI in detecting threats, a capability that becomes increasingly vital as synthetic and multimodal frauds become more prevalent.Advanced AI-driven threat detection acts as a key weapon in the fight against fraud, holding greater significance as the rise of synthetic and multimodal attacks signs.Recent banking cybersecurity research has highlighted the critical importance of advanced AI threat detection in bolstering fraud prevention, particularly as synthetic and multimodal frauds gain prominence.

Organizational Resilience Theory also asserts that the capacity to predict and identify disturbances plays an important role in an organization's capacity to: withstand, respond, adapt, and recover from adverse events. By effectively identifying threats that use the deep fake technology, banks can implement enhanced verification, blocking on the transactions, account protection, investigation of the incidents, and recovery procedures, before a threat escalates. Hence, the impact of DTDE should be translated into tangible fraud-prevention and cybersecurity outcomes, which should then provide theoretical justifications to a possible positive association between DTDE and DBFCR.

H8: Deepfake Threat Detection Effectiveness (DTDE) positively mediates the relationship between AI-Driven Multimodal Deepfake Detection Capability (AMDDFC) and Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

While AMDDFC equips banks with advanced cross-modal analytical tools, it is important to note that these tools are not sufficient to make banks cybersecurity resilient on their own. Partly, the degree of the organization's usefulness should be based on whether they actually help the bank to better detect threats that leverage deep fake technology.Some of that organizational value should derive from whether they actually assist the bank in performing that detection mission. To better identify and highlight inconsistencies between the audio, video, image, biometrics, and contextual evidence, multimodal analysis can be considered.Multimodal analysis may be employed to combine complementary signals with the goal of identifying inconsistencies between the audio, video, images, biometrics, and contextual information, thereby enhancing the identification of sophisticated synthetic fraud. This enhanced capability in threat identification can then help banks to take early fraud control steps.

This type of mediation has combined RBV with Organizational Resilience Theory. From the perspective of RBV the capability of an organization to develop effective action on anticipation and response in a way that creates resilient outcomes is of high value, while from the perspective of resilient theory, technology resources need to be transformed into effective anticipation and response processes before they produce outcomes of resilience; this is why the organization's capability for anticipatory and responding actions may be a valuable capability. DTDE is the mode of operation by which the capability of an AI multimodal becomes useful to combat fraud and to make cybersecurity more resilient. Consequently, the higher the AMDDFC, the higher the DBFCR, and DBFCR would be stronger.As a result, with the higher level of AMDDFC, the higher the level of DTDE which would lead to higher DBFCR.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

H9: Deepfake Threat Detection Effectiveness (DTDE) positively mediates the relationship between AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) and Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

ARTDDC gives banks the ability to conduct a quick analysis of possible deepfakes, but, importantly, its role in organisation resilience should rely significantly on the efficacy of such rapid analysis to detect real threats. Real-time analytics enables the ability to evaluate suspicious identity interactions and synthetic content at the point of interaction, prior to completing a fraudulent process and improve opportunity for prevention (Sarcea, 2024). Therefore, ARTDDC's first priority should be to improve DTDE by ensuring that the identification of deepfakes – a precursor to potential threats – occurs with greater timeliness, with better identification leading to better fraud prevention and cyber-response results.

This indirect relationship is theoretically supported by Dynamic Capabilities Theory, which explains how organizational value is created when an organization detects threats quickly and transforms its findings into appropriate actions. Schreiber & Schreiber, (2024) explains the theory of Dynamic Capabilities in which organizational value is created in doing a threat sensing activity to be followed by values in the form of transformation and transposing the sensed threat into suitable actions. DTDE is the single step between the continuous polling and sensing and the deliverable of helpful security intelligence. The threat once identified can trigger controls that can restrict the exposure to fraud and allow banks to continue to operate in safety. Therefore, DTDE is predicted to be a mediator between the relationship of ARTDDC and DBFCR.

H10: Deepfake Threat Detection Effectiveness (DTDE) positively mediates the relationship between Explainable AI-Based Deepfake Detection Capability (XAIDDC) and Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

While it can enhance the transparency of decisions regarding a deepfake, its significance to the cybersecurity field may become more apparent when the outcome is comprehensible, allowing a bank to better detect and validate threats. Explainable evidence can help cybersecurity teams interpret suspicious or unusual features, assess the validity of AI-generated alerts, and sort out legitimate threats from reported ones but not necessarily correct. Recent studies highlight that, as financial AI moves into a new era where trust and transparency are paramount, explainability becomes just as vital as the ability to detect frauds and handle risks. XAIDDC should thus need to play a role in enhancing DTDE, which is characterized by the ability to take an opaque prediction and make it a comprehensible threat threat that can be analysed. Therefore, XAIDDC should help enhance DTDE by transforming opaque predictions into understandable threat intelligence that could be analysed.

The mediation mechanism is explained as part of OIPT and of Organizational Resilience Theory. Resilience theory points to the fact that helpful threat knowledge allows for looking ahead, responding, adapting, and recovering and OIPT suggests that helpful knowledge helps improve organizational processing and decision making. Therefore, explainability will be expected to boost the resilience, at least in increasing the identification and validation of deepfake threats. DTDE thus acts as the linkage mechanism between XAIDDC and the overall fraud and Cyber security outcomes, underscores a positive indirect relationship between XAIDDC and the DBFCR (Rafy, 2024).

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

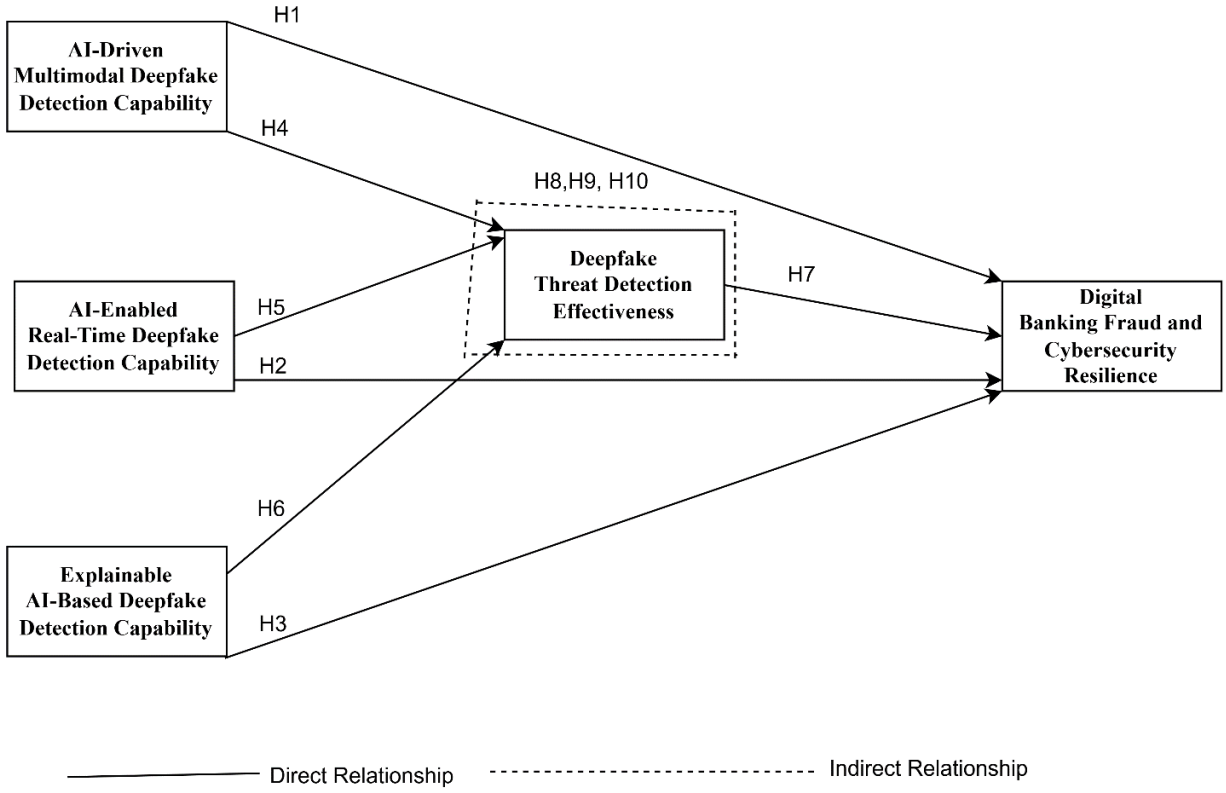


Figure 1: Proposed Research Model

RESEARCH METHODOLOGY

Research Design

This is a quantitative, explanatory and cross-sectional research design that leverages empirical methods to understand the relationships between AI-Driven Multimodal Deepfake Detection Capability (AI-DMMDFDC), AI-Enabled Real-Time Deepfake Detection Capability (AI-ERTDFC), Explainable AI-Based Deepfake Detection Capability (AI-ExplainableDFDC), Deepfake Threat Detection Effectiveness (DTDE), and Digital Banking Fraud and Cybersecurity Resilience (DBFCR) in the U.S. digital banking systems. Structured survey with multiple reflective indicators on all the latent constructs were used with primary data collection. Based on the abovementioned aspects, in order to estimate multiple latent_variable relationships, test mediation effects and analyse the explanatory and predictive abilities of the theoretically developed model, the researcher decided to use Partial Least Squares Structural Equation Modeling (PLS-SEM) using the SmartPLS 4 software. Recent empirical research has also used quantitative cross sectional designs with PLS-SEM for complex organizational and technology related models. This analysis was conducted in two stages, in the first the measurement model was analyzed based on the indicator loading, internal consistency reliability, convergent validity, discriminant validity, and collinearity diagnosis, and in the second stage the analysis of the structural model was carried out based on path coefficients, VIF, coefficient of determination (R^2), effect size (f^2), direct and indirect effects, and predictive relevance. Significance tests and confidence intervals were computed on the hypothesized direct and mediating relationships using 5,000 bootstraps, which allowed the evaluation of whether the hypothesized direct and mediating relationships were statistically significant in the study and provide the proper confidence intervals between these hypotheses.

Measurement Instruments:

In order to maintain all five constructs proposed in the research model introduced a measurement instrument containing five multi-item reflective constructs: AI-Driven Multimodal Deepfake Detection Capability (AMDDFC), AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC), Explainable AI-Based Deepfake Detection Capability (XAIDDC), Deepfake Threat Detection Effectiveness (DTDE), and Digital Banking Fraud and

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Cybersecurity Resilience (DBFCR). Six items were used to measure each of six constructs giving a total of 30 measurement items. The fact that there were no previous scales directly validated for organizational-level scales for all five constructs of deepfakes makes authors decide not to use them verbatim from the literature on multimodal deepfake detection, real-time fraud detection, explainable AI, synthetic-identity threats, and financial cybersecurity, but rather to adapt and place them appropriately within the literature (Ali et al., 2022). Items were all in the past tense to align with the retrospective research design, and were assessed on a five-point Likert scale with 5 indicating strongly agree, and 1 indicating strongly disagree. It therefore measured the extent of U.S. digital banking organizations' developing capabilities in the areas most relevant to the issue at hand – the ability to detect deep fakes and to increase fraud and cybersecurity resilience.

Table 1: Measurement Items for Study Constructs

Variable	Code	Measurement Item
AI-Driven Multimodal Deepfake Detection Capability	AMDDFC1	The bank had used AI to jointly analyze facial/video, voice/audio, image, and other digital signals for deepfake detection.
	AMDDFC2	The bank had integrated information from multiple modalities to identify manipulated or synthetic digital content.
	AMDDFC3	The bank's AI systems had compared information across different modalities to identify inconsistencies associated with deepfakes.
	AMDDFC4	The bank had combined multiple sources of digital evidence when assessing suspected deepfake-based impersonation.
	AMDDFC5	The bank's multimodal AI capability had enabled it to identify deepfake threats that could have been overlooked by single-modality detection.
	AMDDFC6	The bank had continuously improved its multimodal AI capability to detect increasingly sophisticated synthetic content.
AI-Enabled Real-Time Deepfake Detection Capability	ARTDDC1	The bank had used AI to detect suspected deepfake activity in real time during digital banking interactions.
	ARTDDC2	The bank had rapidly analyzed suspicious synthetic content during digital customer onboarding.
	ARTDDC3	The bank had detected deepfake-based impersonation during digital identity authentication processes.
	ARTDDC4	The bank had monitored transactions and customer interactions for deepfake-related threats as they occurred.
	ARTDDC5	The bank's AI systems had generated timely alerts when potential deepfake manipulation was detected.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Variable	Code	Measurement Item
	ARTDDC6	The bank had rapidly adapted its deepfake detection processes in response to emerging forms of AI-enabled fraud.
Explainable AI-Based Deepfake Detection Capability	XAIDDC1	The bank's AI systems had provided understandable explanations for why digital content was classified as genuine or manipulated.
	XAIDDC2	Cybersecurity personnel had been able to understand the main factors underlying AI-generated deepfake alerts.
	XAIDDC3	The bank's AI systems had provided interpretable evidence supporting deepfake detection decisions.
	XAIDDC4	The bank had enabled human security personnel to review and validate AI-generated deepfake classifications.
	XAIDDC5	The bank had maintained transparency regarding how AI-supported deepfake detection decisions were generated.
	XAIDDC6	Explainable AI outputs had enabled security personnel to determine appropriate responses to suspected deepfake threats.
Deepfake Threat Detection Effectiveness	DTDE1	The bank had effectively identified deepfake-based customer or employee impersonation attempts.
	DTDE2	The bank had effectively detected AI-generated or cloned voices used in fraudulent interactions.
	DTDE3	The bank had effectively identified manipulated facial or video evidence used in digital banking activities.
	DTDE4	The bank had effectively detected synthetic identities used to obtain unauthorized access to banking services.
	DTDE5	The bank had effectively distinguished genuine digital interactions from deepfake-enabled fraudulent interactions.
	DTDE6	The bank had effectively identified emerging deepfake-enabled fraud attempts before they caused significant security consequences.
Digital Banking Fraud and Cybersecurity Resilience	DBFCR1	The bank had effectively prevented financial fraud arising from deepfake-enabled attacks.
	DBFCR2	The bank had maintained secure digital banking operations when confronted with deepfake-related cybersecurity threats.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Variable	Code	Measurement Item
	DBFCR3	The bank had effectively contained the impact of deepfake-enabled fraud and cybersecurity incidents.
	DBFCR4	The bank had responded rapidly to cybersecurity incidents involving synthetic identities or manipulated digital content.
	DBFCR5	The bank had effectively recovered from disruptions caused by deepfake-enabled fraud or cybersecurity incidents.
	DBFCR6	The bank's cybersecurity arrangements had adapted effectively to evolving deepfake and AI-enabled fraud threats.

The wording was carefully formulated to be organization level and capability centre as it was aimed at assessing the theoretical construct of the study and not respondents' personal ability to detect deepfakes. The multimodal integration and cross-modal reasoning domains, which demanded effective deepfake detection, were previously highlighted in the recent literature; for synthetic and financial risks, the most relevant content was the real-time deepfake detection and risk metric; and explainability lay at the core of the recent literature in dealing with the interpretation and operationalization of decisions supported by AI tools for security purposes. DTDE, as a result, reflected the effectiveness of realized threat identification and the DBFCR reflected the wider organizational products of prevention, continuity, containment, response, recovery, and adaptation.

Population and Sampling

Professionals with roles or expertise that fall under the banking and digital financial-service space in the U.S. who have a role or knowledge concerning digital banking, cybersecurity, fraud prevention, information technology (IT), risk management, digital identity verification, compliance or AI driven security systems were included in the study. Purposive sampling was the most suitable method as the study aimed to engage respondents who had relevant knowledge and experience to provide informed assessments of the quality of (AI) Deepfake Detection Capability (DDC) and (AI) Explainable AI-Based DDC (XAIDDC), a RealTime Deepfake Detection Capability (RTDDC), a Deepfake Threat Detection Effectiveness (DTDE), and a Digital Banking Fraud and Cybersecurity Resilience (DBFCR). In recent studies on cybersecurity, banking, and technology in the quantitative approach, purposive sampling strategy was adopted for certain quantitative studies for which specialized or relevant experience was required to take part in the study (Bhardwaj & Goundar, 2024). The final sample consisted of 510 respondents who were valid for the respective professional position in a banking organization and a representation of the digital banking operations in the U.S. banking sector. This value was deemed sufficient for PLS-SEM due to the fact that it includes a relatively large number of samples for stable estimation of the measurement model, structural model (including multiple direct paths) and mediation models (including the three proposed mediation effects) for different types of models; recent research on PLS-SEM has also stated that in addition to minimum sample size rules, samples with a relatively large number of samples are very important for stable estimation of the measurement model, structural model (including multiple direct paths) and the mediation model (including the three proposed mediation effects) in various types of model (Kazimierczak et al., 2024). For this reason, using a purposive respondent selection method between the beginning of the research and a final sample of 510 respondents, was sufficient to establish the empirical basis for the investigation of the proposed AI deepfake-detection and cybersecurity-resilience model.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Data Collection Method

The primary data were gathered by distributing a structured and self-administered online questionnaire to the professionals working in U.S. banking and digital financial-service organisations following the same procedure as used in various recent quantitative banking, cybersecurity, and technology studies to collect standardized information for PLS-SEM analysis (Pleshakova et al., 2024). The questionnaire consisted of two parts: one asking for the demographic and professional data of the respondents and the other asking the 30 measurement items that represent the capabilities of AI-Driven Multimodal Deepfake Detection Capability (AMDDFC), AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC), Explainable AI-Based Deepfake Detection Capability (XAIDDC), Deepfake Threat Detection Effectiveness (DTDE), and Digital Banking Fraud and Cybersecurity Resilience (DBFCR). The items on the constructs were measured on a 5-point Likert scale, from 1 = strongly disagree to 5 = strongly agree. To keep the time focused on the development of the interview guide as well as ensuring that participants had the required knowledge or experience in digital banking, cyber security and fraud prevention, risk management or information technologies, and AI-powered security and digital identities, a purposive data collection approach was used, similar to recent studies that targeted knowledgeable respondents for their specific research in cyber security and financial technology. No personally identifying information was requested and participants were informed about academic purpose of the study and the voluntary participation, confidentiality and anonymity of the responses before completing the questionnaire. The questionnaires returned were then checked for completeness and quality of responses so that finally 510 valid questionnaire responses could be kept, coded and processed for measurement and analysis of structural model using PLS SEM.

Respondents' Demographic Profile

The respondents (n = 510) were quite diverse in the demographic profile across gender, age, education, organizational position, professional experience and industry experience. In terms of gender, 326 respondents (63.9%) were male and 184 (36.1%) were female. Regarding age, 38 respondents (7.5%) were below 25 years, 186 (36.5%) were aged 25–34 years, 164 (32.2%) were aged 35–44 years, 91 (17.8%) were aged 45–54 years, and 31 (6.1%) were aged 55 years or above, indicating that most participants were within the professionally active 25–44 age range. As far as educational background is concerned, 142 (27.8%) responded with a bachelor's degree, 248 (48.6%) with a master's degrees, 103 (20.2%) with MPhil/PhD and 17 (3.3%) with other educational background, reflecting that the group of respondents was very educated. As far as position within the organization, 128 respondents (25.1%) work in managerial roles, 104 (20.4%) are senior executives, 139 (27.3%) are business/data analysts, 73 (14.3%) are sustainability professionals and 66 (12.9%) work in other roles. Professional experience was also very spread out, with 82 (16.1%) having < 5 years of experience, 174 (34.1%) having 5–10 years of experience, 146 (28.6%) having 11–15 years of experience, and 108 (21.2%) having ≥ 16 years of experience. Lastly, industry distribution revealed that 148 respondents (29.0%) were working in the manufacturing industry whereas 91 (17.8%) were working in the energy industry, 106 (20.8%) in technology industry, 74 (14.5%) were working in agriculture industry and also 91 (17.8%) respondents were working in other industries. The demographic profile also showed that the sample consisted of individuals with diverse backgrounds in terms of education, occupation, and experience, providing a solid empirical foundation for the subsequent PLS-SEM analysis; different demographic characteristics, including gender, age, education, occupation, and experience, have also been described in recent studies on digital banking and cybersecurity (Kaushik et al., 2024).

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

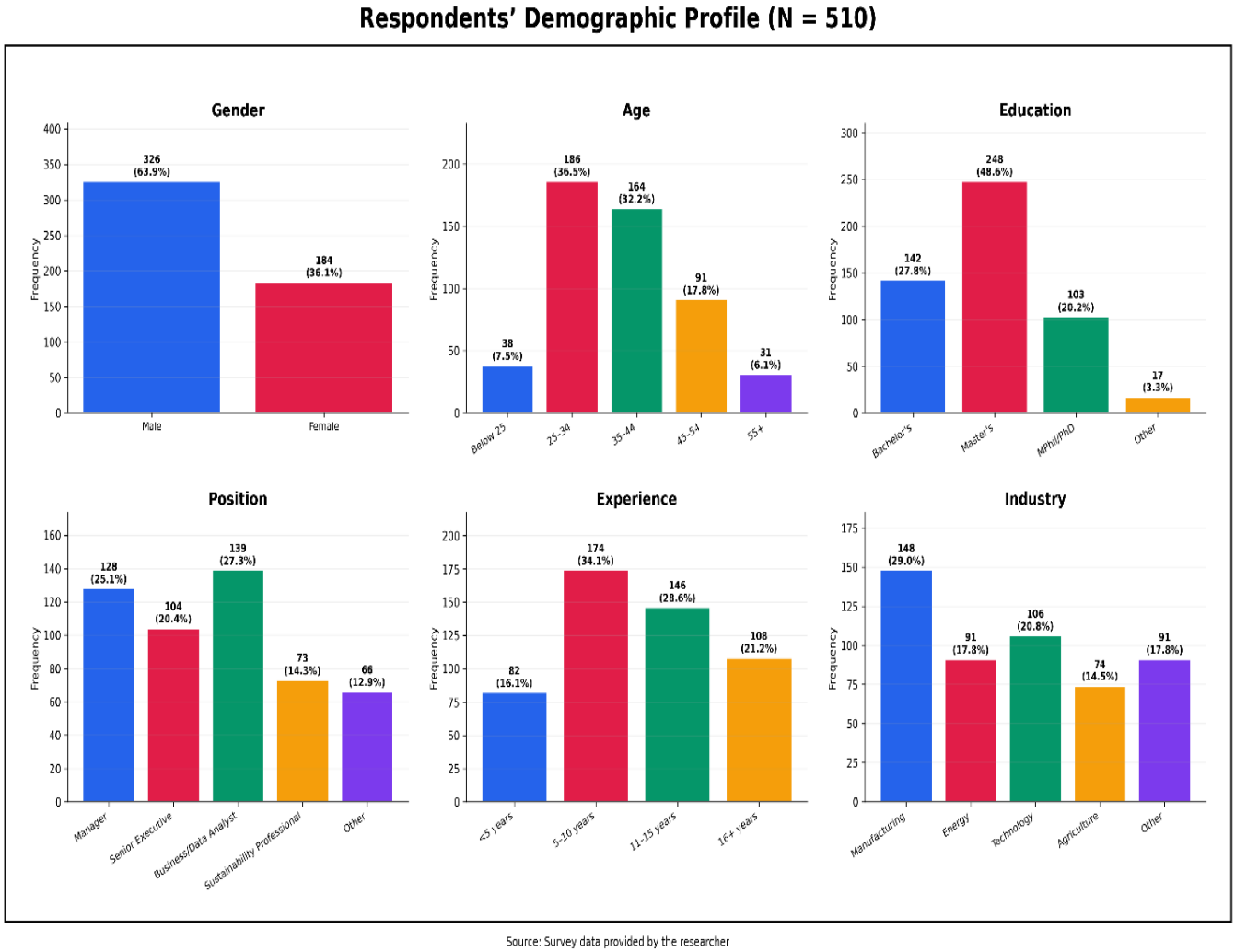


Figure 2: Respondents Profile

Data Analysis Technique

All the collected data were analyzed using the Partial Least Squares Structural Equation Modeling (PLS-SEM) via the Software of SmartPLS 4, which would allow the direct and indirect linkages between the five latent variables and measurement evaluation. Analytical stages were followed:Two stages of analysis were done: First, the indicator reliability, internal consistency, convergent and discriminant validity as well as problematic multicollinearity were determined in terms of the reflective measurement model by employing indicator loadings, Cronbach's alpha, composite reliability (CR), average variance extracted (AVE) and Variance inflation factor (VIF) and heterotrait–monotrait ratio (HTMT) respectively. First, standardized path coefficients (β), coefficients of determination (R^2), effect sizes (f^2), predictive relevance factor (coefficient π) and direct and specific indirect effects were calculated for the whole structural model to measure the explanatory and predictive performance of the proposed model. The statistical significance of the hypothesized relationships were examined using a nonparametric bootstrapping procedure with 5,000 resamples, and t-values, p-values and 95% confidence intervals showed that the proposed effects were also statistically supported as per the recent applications of PLS-SEM which were applied by Pashentsev & Bazarkina, (2023b) respectively. Subsequently mediation analysis was conducted using bootstrapped specific indirect effects, which overall supported the study's measurement and structural models, confirm their links between the constructs, and quantify the impact of AI-Driven Multimodal Deepfake Detection Capability (AMDDFC), AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC), and Explainable AI-Based Deepfake Detection Capability (XAIDDC) on Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Assessment of Measurement Model

This was followed by the assessment of the measurement model in PLS-SEM in Smart PLS 4 to ensure the reliability and validity of the five reflective measures: AI-Driven Multimodal Deepfake Detection Capability (AI-DMDDFC); AI Enabled Real-Time Deepfake Detection Capability (AI-ERTDDC); Explainable AI Based Deepfake Detection Capability (Explainable AI (E) DDFC); Deepfake Threat Detection Effectiveness (DTDE); and Digital banking Fraud and Cybersecurity Resilience (DBFCR). Dropping indicators for reliability of the indicators was done based on outer loadings of ~ 0.708 and less for values higher than 0.708 were regarded as desirable, checking of variance inflation factor (VIF) at the indicator level were then carried out to obtain indicators that have potential collinearity problems. Internal consistency reliability was determined by Cronbach's alpha regression coefficient ranging between 0.70 and above considered satisfactory and composite reliability (CR), with values of at least 0.50 that means that each construct explains over half the variance of its indicators was considered satisfactory, convergent validity was determined by the average variance extracted (AVE) regression coefficient of which values of 0.50 or above were considered satisfactory (Kaushal et al., 2024). The heterotrait – monotrait ratio (HTMT) was then checked to assess discriminant validity by assessing differences between constructs, in order to detect various constructs from each other. The assessment thus prioritized whether the indicators were suitable for their intended constructs and reliability, convergent validity, discriminant validity and collinearity were addressed prior to the assessment of the structural model and the direct and mediating hypotheses measured.

Common method bias assessment showed that the common method bias was not a serious problem in this study. After Principal Axis Factoring the first factor had an initial eigenvalue of 12.143 to account for 40.478% of the total variance, and produced 39.693 % of the variance after extraction. In both instances the variance in the measurement items in the first factor was under 50% which most studies suggest does not make common method bias a significant problem (Ali et al., 2023). Further, there had been some variance spreading out across multiple factors: the first 3 reported extracted factors together accounted for 70.349% of the variance, and the five reported extracted factors each together accounted for 77.227% of the variance that had been extracted, further indicating that such responses were not dominated by any single general methodological factor. To the best of our knowledge, the present study is the first to use Harman's single-factor procedure as a preliminary assessment of common method bias in questionnaire-based PLS-SEM studies in parallel with the recent ones.To date, as far as we know, no similar studies have used Harman's single-factor procedure as a preliminary test for the presence of common method bias in questionnaire-based PLS-SEM research, besides the recent studies like our study. Thus, the results indicated that there was little concern about occurrence of common method bias as it had not materially affected the relationships among the study constructs; therefore, the results could be accepted for future measurement and structural model analyses.

From the results in Table 3, it was obtained that the indicator reliability of the measurement model was quite high, internal consistency reliability of the measurement model was good, and the convergent validity of the measurement model was good as well as indicator collinearity of the measurement model showed acceptable values. The factor loading values for all the indicators were above the recommended 0.708 suggesting that all the indicators were adequate of capturing the latent constructs of DTDE1 (from 0.850) to XAIDDC2 (from 0.927). Cronbach's alpha values ranged between 0.942 (DTDE) and 0.965 (XAIDDC) and composite reliability values ranged from 0.954 to 0.971, which was much higher than the widely accepted minimum reliability of 0.70. Likewise, the AVE for each item has ranged from 0.775 for DTDE to 0.850 for XAIDDC much higher than the recommended AVE value (0.50); this ensured good convergent validity. Specifically, ARTDDC had reported $\alpha = 0.958$, CR = 0.966, and AVE = 0.827; AMMDFC had reported $\alpha = 0.948$, CR = 0.958, and AVE = 0.793; DTDE had recorded $\alpha = 0.942$, CR = 0.954, and AVE = 0.775; DBFCR had produced $\alpha = 0.949$, CR = 0.959, and AVE = 0.796; and XAIDDC had demonstrated the highest reliability and convergent validity with $\alpha = 0.965$, CR = 0.971, and AVE = 0.850. Moreover, most of the VIF varied from 2,643 to 4,867, which did not show problematic indicator-level multicollinearity as is usually considered. In general, the results had shown that all 30 indicators were valid and reliable; and the five constructs were valid and had acceptable psychometric properties for further testing of the discriminant-validity and the structural model.

Concerning the discriminatory validity of the five constructs, the results shown in Table 4 suggested that adequate discriminatory values were achieved among the five constructs based on the Heterotrait-Monotrait (HTMT) ratio.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The highest was found between Digital Banking Fraud and Cybersecurity Resilience (DBFCR) at 0.763; followed by Deepfake Threat Detection Effectiveness (DTDE) and ARTDDC at 0.546 and 0.476 respectively; followed by XAIDDC (0.468), AMMDFC (0.447), AMMDFC (0.437), XAIDDC (0.423) and DBFCR (0.423). The other relationships were significantly lower with smaller HTMT values such as AMMDFC–XAIDDC (0.034), ARTDDC–XAIDDC (0.058) and AMMDFC–ARTDDC (0.026). Most importantly, none of the relationship between the latent constructs showed any problematic conceptual overlap as all of the HTMT were below the benchmark value of 0.85 in modern PLS-SEM studies. The comparatively larger HTMT value between DTDE and DBFCR (0.763) indicated these two constructs had meaningful conceptual association, however it was not so close to 0.85 (conservative threshold) or 0.9 (more liberal threshold). Thus, the findings confirmed the adequate discriminant validity between the different constructs (AMMDFC; ARTDDC; DBFCR; DTDE; XAIDDC) justifying their use in subsequent analyses of the model and hypotheses considered.

Results of the effect size are presented in the Table 5 and these results showed that the latencies have exerted different effect when related with the endogenous constructs. From a commonly adopted perspective on the size of the effects of f^2, which is around 0.02, 0.15 and 0.35 respectively for small, medium, and large effects, the contribution of AMMDFC was comparatively stronger in enhancing the DTDE (f^2 = 0.344), than in improving the DBFCR (f^2 = 0.148). ARTDDC had shown to have a medium effect on DBFCR (f^2 = 0.235) and a large effect on DTDE (f^2 = 0.395) which was the biggest effect in the model. The DTDE had levelled a moderate influence on DBFCR (R2 = 0.162), and played a significant role in accounting for the banking fraud and cybersecurity resilience. XAIDDC had demonstrated a small f^2 value for DBFCR (f^2 = 0.098) and a larger f^2 value for DTDE (f^2 = 0.369) as compared to DBFCR. This classification corresponds with the recent PLS-SEM-based application that used 0.02, 0.15 and 0.35 as thresholds separating small, medium and large effects, respectively. Overall, the results indicated that the substantive role of DTDE in the proposed model strengthened by the contribution of ARTDDC and XAIDDC however, in the case of DBFCR it was seen that the contribution of AI techniques for the detection of deepfake images were relatively insubstantial on a small to medium scale.

Table 6 presented the results of the model-fit that showed satisfactory explanatory capability, predictive relevance, and approximate model fit of the structural based on the results obtained. The R2 for Digital Banking Fraud and Cybersecurity Resilience (DBFCR) was 0.640, with an adjusted R2 or 63.7% that showed that AMMDFC along with ARTDDC, XAIDDC, and DTDE accounted for 64.0% of the variance in DBFCR. Likewise, DTDE attained an R^2 of 0.542 and an adjusted R2 of 0.539 indicating that the three AI-based deep fake detection mechanisms had accounted for 54.2% of its variance. Prior to the PLS-SEM application, the Q^2predict values were positive for both constructs of endogenous variables, namely for DBFCR (+0.536) and DTDE (+0.576), as is common in previous PLS-SEM applications, where the positive Q^2predict values have been considered as evidence for out-of-sample predictive relevance (Folds, 2022). The prediction-error measures of RMSE = 0.685 and MAE = 0.533 for DBFCR and RMSE = 0.654 and MAE = 0.529 for DTDE quantified prediction errors, but only a comparison with an appropriate benchmark model can demonstrate a superior prediction performance, not looking at them separately. Additionally, the SRMR value for both saturated model and estimated model were 0.026, which are fairly low compared to the widely accepted criterion of 0.08 (Kovací, 2023) indicating satisfactory approximate model fit. Moreover, the value obtained for the NFI was high, 0.961 for both models, indicating high relative fit values. Overall, the results had indicated that the proposed model accounted for significant variance in DBFCR and DTDE, was relevant with a positive predictive role and, furthermore, was able to provide satisfactory approximate fit for its further interpretation of the structural connections and results obtained from a test of the hypotheses.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

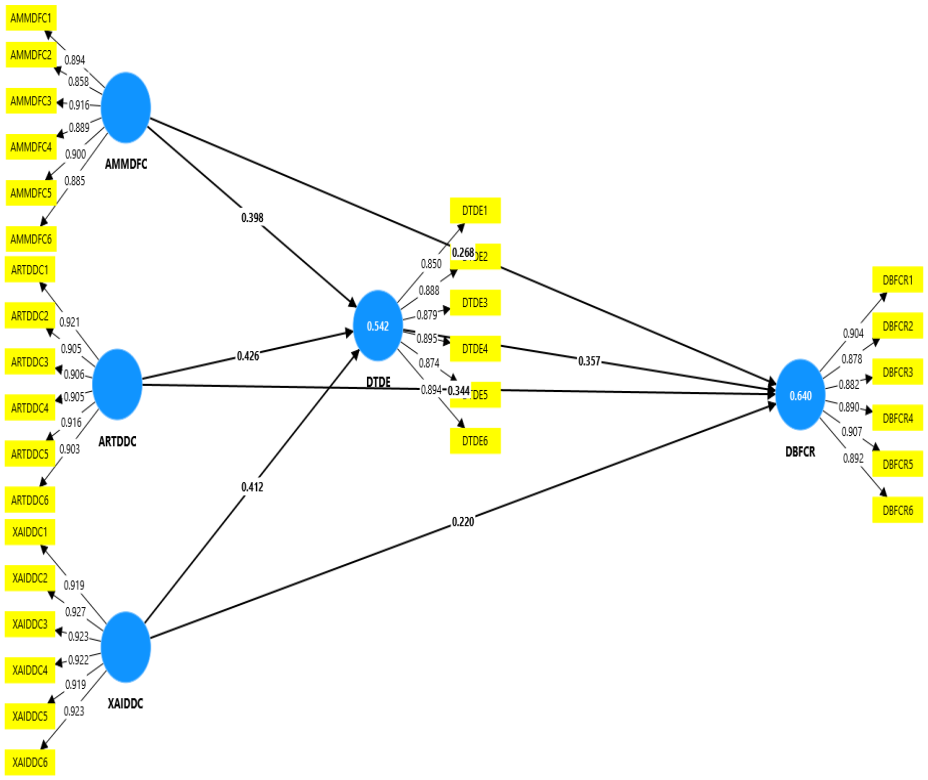


Figure 3: SEM with Factor Loadings

Table 2: Common Method Bias

Total Variance Explained						
Factor	**Initial Eigenvalues**			**Extraction Sums of Squared Loadings**		
	Total	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	12.143	40.478	40.478	11.908	39.693	39.693
2	4.819	16.064	56.542	4.620	15.401	55.094
3	4.797	15.989	72.532	4.576	15.255	70.349
4	1.338	4.460	76.991	1.087	3.623	73.972
5	1.205	4.018	81.009	.976	3.255	77.227
Extraction Method: Principal Axis Factoring.						

Table 3: Reliability, Validity and Variance Inflation Factor (VIF)

Construct/Items	**VIF values**	**Factor Loadings**	**Cronbach's Alpha**	**Composite Reliability**	**Average Variance Extracted (AVE)**
Predictive Analytics Capability			0.958	0.966	0.827
ARTDDC1	4.519	0.921			
ARTDDC2	3.887	0.905			

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

ARTDDC3	3.949	0.906			
ARTDDC4	3.905	0.905			
ARTDDC5	4.333	0.916			
ARTDDC6	3.918	0.903			
AI-Driven Multimodal Deepfake Detection Capability			0.948	0.958	0.793
AMMDFC1	3.593	0.894			
AMMDFC2	2.809	0.858			
AMMDFC3	4.292	0.916			
AMMDFC4	3.386	0.889			
AMMDFC5	3.736	0.900			
AMMDFC6	3.305	0.885			
Deepfake Threat Detection Effectiveness			0.942	0.954	0.775
DTDE1	2.643	0.850			
DTDE2	3.311	0.888			
DTDE3	3.118	0.879			
DTDE4	3.519	0.895			
DTDE5	3.024	0.874			
DTDE6	3.486	0.894			
Digital Banking Fraud and Cybersecurity Resilience			0.949	0.959	0.796
DBFCR1	3.873	0.904			
DBFCR2	3.131	0.878			
DBFCR3	3.207	0.882			
DBFCR4	3.421	0.890			
DBFCR5	3.973	0.907			
DBFCR6	3.410	0.892			
Explainable AI-Based Deepfake Detection Capability			0.965	0.971	0.850
XAIDDC1	4.568	0.919			
XAIDDC2	4.867	0.927			
XAIDDC3	4.792	0.923			
XAIDDC4	4.553	0.922			
XAIDDC5	4.594	0.919			
XAIDDC6	4.778	0.923			

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 4: Discriminate Validity Heterotrait-Monotrait (HTMT) Ratio

	AMMDFC	ARTDDC	DBFCR	DTDE	XAIDDC
AMMDFC					
ARTDDC	0.026				
DBFCR	0.447	0.546			
DTDE	0.437	0.476	0.763		
XAIDDC	0.034	0.058	0.423	0.468	

Table 5: Effect Size of Latent Variables

	f-square	Effect Size
AMMDFC -> DBFCR	0.148	Small
AMMDFC -> DTDE	0.344	Medium
ARTDDC -> DBFCR	0.235	Medium
ARTDDC -> DTDE	0.395	Large
DTDE -> DBFCR	0.162	Medium
XAIDDC -> DBFCR	0.098	Small
XAIDDC -> DTDE	0.369	Large

Table 6: Model Fit

	R-sq	Adj R-sq	Q^2predict	RMSE	MAE	Saturated Model	Estimated Model
DBFCR	0.640	0.637	0.536	0.685	0.533		
DTDE	0.542	0.539	0.576	0.654	0.529		
SRMR						0.026	0.026
NFI						0.961	0.961

Hypotheses Testing

H1: AI-Driven Multimodal Deepfake Detection Capability (AMMDFC) positively influenced Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

However, the present results had confirmed that H1, AMMDFC, had a statistically significant, positive effect on DBFCR; a t value of 8.914 with a p value of < 0.001. The 95% bootstrapped confidence interval had not contained zero and ranged from 0.209 to 0.326 further reinforcing the significance of the relationship. This discovery suggested that an increase in ability to integrate and analyze more than one modality relevant to deepfakes had been linked to increased fraud prevention and cyber security strength in digital banking. Thus, it had been accepted that H1 was true. The interpretation was consistent with the PLS-SEM interpretation criteria set above, which is a structural relationship that is statistically significant based on t-value above 1.96, p value < 0.05, and confidence interval value not included in 0.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

H2: AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) positively influenced Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

The findings had shown that ARTDDC had acted positively and statistically significantly with regards to DBFCR, with the t-value of 11.533 the p-value of less than 0.001. The 95% CONFIDENCE_INTERVAL was now longer, banging between 0.286 and 0.403 with no negative numbers. Such a finding was said to indicate banking firms' ability to adequately identify threats stemming from deepfakes quickly during digital interactions had been linked to higher capability to prevent, tolerate and react to fraud or cybersafety events, respectively. This meant that H2 was accepted.

H3: Explainable AI-Based Deepfake Detection Capability (XAIDDC) positively influenced Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

The findings revealed that H3 has been supported: As for the effects of XAIDDC while thеκ value was 0.220, t-value was 7.477 and p-value was < 0.001, which indicates that the XAIDDC had caused a positive and significant effect on DBFCR. The 95% confidence interval for its had been (0.161, 0.278) which did not contain zero and hence supported the statistical significance of the relationship. The findings showed that explainable and comprehensible AI-provided evidence for deepfake detection was related with enhanced digital business fraud prevention and Cyber-resilience. As such, the proposal had been accepted for H3.

H4: AI-Driven Multimodal Deepfake Detection Capability (AMMDFC) positively influenced Deepfake Threat Detection Effectiveness (DTDE).

H4 had been supported by the findings as AMMDFC had shown a positive relationship with DTDE being statistically significant ($\beta = 0.398$, $t = 14.297$, $p < 0.001$). The interval associated with the reported 95 percent confidence interval was (0.293, 0.423), which did not contain 0.The interval (0.293, 0.423) was the 95 percent confidence interval reported and contained no 0's. The results showed that the collective intelligence of facial/video, voice/audio, image and other digital signals – the combination of which has never been possible to achieve until now – were positively correlated with the effectiveness of banks' identification of deepfake impersonation, synthetic identity, manipulated media and associated fraud. So, it was accepted that H4 was the case.

H5: AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) positively influenced Deepfake Threat Detection Effectiveness (DTDE).

Finally, the results were supportive of H5 as evidenced by elevated and statistically significant beta (0.426), t (14.735) and p value (<0.001) between ARTDDC and DTDE. The 95% confidence interval in previous years had been between 0.342 and 0.452 and it was positive. The results showed that early detection and monitoring of deepfakes in the onboarding, authentication, transactions and customer interaction processes could result in the potential for increased effectiveness in detecting deepfakes within their organization. So, H5 was accepted in this case.

H6: Explainable AI-Based Deepfake Detection Capability (XAIDDC) positively influenced Deepfake Threat Detection Effectiveness (DTDE).

The results had confirmed H6 which had been positively and significantly influenced DTDE ($\beta = 0.412$, $t = 14.886$, $p < 0.001$). The bootstrapped 95% confidence interval for the parameter had not contained zero but has been found to lie within 0.368 and 0.481. This finding led them to believe that making the explanations and evidence provided by the AI easy to understand and interpret, has been correlated with higher effectiveness in detecting threats in the image. The evidence from the statistics, therefore, had been in favour of accepting H6.

H7: Deepfake Threat Detection Effectiveness (DTDE) positively influenced Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

The results had confirmed the hypothesis as the DTDE had been positively correlated with DBFCR with strong t value ($t = 10.692$ and $p < 0.001$). The reported 95% confidence interval from (0.357, 0.466) was still positive and consistent with other commonly-used bootstrapping criteria, indicating statistical significance. The discovery showed that higher effectiveness in preventing deepfakes, voice cloning, manipulated contents, synthetic identities, and almost any other type of fraud that leverages AI has been linked to cyber fraudR and digital banking resilience. As such, H7 had been positively accepted and empirically supported as an aspect of the model's explanation of the phenomenon.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 7: Hypotheses Testing Direct Relationship (Bootstrapping)

No.	Paths	β- Values	t- Values	p- Values	2.5%	97.5%	Results
H1	AMMDFC→DBFCR	0.268	8.914	0.000	0.209	0.326	Accepted
H2	ARTDDC→ DBFCR	0.344	11.533	0.000	0.286	0.403	Accepted
H3	XAIDDC→ DBFCR	0.220	7.477	0.000	0.161	0.278	Accepted
H4	AMMDFC→ DTDE	0.398	14.297	0.000	0.293	0.423	Accepted
H5	ARTDDC→ DTDE	0.426	14.735	0.000	0.342	0.452	Accepted
H6	XAIDDC→ DTDE	0.412	14.886	0.000	0.368	0.481	Accepted
H7	DTDE→DBFCR	0.357	10.692	0.000	0.357	0.466	Accepted

H8: Deepfake Threat Detection Effectiveness (DTDE) positively mediated the relationship between AI-Driven Multimodal Deepfake Detection Capability (AMMDFC) and Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

The results had validated H8, which was in terms of indirect effect of AMMDFC on DBFCR through DTDE got positive result and statistically significant ($\beta = 0.142$, $t = 8.684$, $p < 0.001$). The direct effect of AMMDFC on DBFCR was also positive ($\beta = 0.268$ when total effect $\beta = 0.410$. The computed VAF value of 34.63% suggests that about a third of the overall effect of AMMDFC on DBFCR was indirect via DTDE. The results had shown partial mediation as the VAF value had been in the range of 20%–80%, which is the acceptable range for partial mediation (PASHENTSEV et al., 2024). Therefore, AMMDFC positively impacted DBFCR, directly and indirectly, by increasing detection capabilities at the bank for deep fake attacks and H8 was accepted.

H9: Deepfake Threat Detection Effectiveness (DTDE) positively mediated the relationship between AI-Enabled Real-Time Deepfake Detection Capability (ARTDDC) and Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

The results had found positive and statistically significant results for the indirect effect of ARTDDC on DBFCR via DTDE ($\beta = 0.152$, $t = 8.766$, $p < 0.001$) that supported hypothesis H9. The direct effect turned out to be positive at $\beta = 0.344$ and the total effect reached $\beta = 0.496$ which showed that the ARTDDC had both direct and mediated effect on DBFCR. However, the VAF value of 30.65% indicated that DTDE was responsible for a significant portion of the overall relationship between the variables under investigation, in the meantime, there was still a direct relationship between the variables. The value of this ranged between 20% and 80%, which are traditionally considered as partial mediation range in PLS-SEM studies. Hence, realtime deepfake detection capability had worked on increasing the efficacy of deepfake threat detection to bolster the resilience of banking fraud and cybersecurity...that is why H9 was accepted.

H10: Deepfake Threat Detection Effectiveness (DTDE) positively mediated the relationship between Explainable AI-Based Deepfake Detection Capability (XAIDDC) and Digital Banking Fraud and Cybersecurity Resilience (DBFCR).

This finding had been congruent with H10 which revealed that DTDE operated indirectly to impact DBFCR via XAIDDC, and had attained both positive and statistically significant β (0.147) ($t = 8.478$) ($p < 0.001$). The direct effect of XAIDDC on the DBFCR was found to be $\beta = 0.220$ while the total effect was $\beta = 0.367$. Based on the VAF value, the DTDE value was 40.05% which meant it had mediated about 40% of the overall relationship between XAIDDC and DBFCR, which was the highest mediation proportion among the three mediation paths listed in Table 8 above. As the value of VAF remained between 20% and 80%, direct effect and indirect effect have been positive, this had been supportive in terms of partial mediation. As a consequence, introducible Artificial Intelligence capability has been directly connected to banking fraud and cyber security resilience, also indirectly, thanks to the better

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

performance in the deepfakes threat detection procedure.H10 was as a consequence accepted, and explainable-AI capability had been directly connected to banking fraud and cyber security resilience also indirectly because of the better performance in deepfakes threat detection procedure.

Table 8: Mediation Type and Effect

No.	Paths	Indirect Effect (β)	t-value	P Value	Direct Effect (β)	Total Effect (β)	VAF (%)	Mediation Result
H8	AMMDFC→DTDE →DBFCR	0.142	8.684	0.000	0.268	0.410	34.63%	Partial Mediation
H9	ARTDDC→ DTDE→DBFCR	0.152	8.766	0.000	0.344	0.496	30.65%	Partial Mediation
H10	XAIDDC→DTDE→ DBFCR	0.147	8.478	0.000	0.220	0.367	40.05%	Partial Mediation

DISCUSSION:

The results showed that AI-powered deepfake detection methods have effectively contributed to the overall improvements in fraud prevention and the wider cybersecurity and resilience efforts of digital banking, both directly, and indirectly via Deepfake Threat Detection Effectiveness (DTDE). All seven of the direct hypotheses were supported and all three of the mediation hypotheses. For Digital Banking Fraud and Cybersecurity Resilience (DBFCR) the model accounted for 64.0% of the variance, for DTDE the model accounted for 54.2% of the variance, and the positive Q^2predict values for both DBFCR (0.536) and DTDE (0.576) showed predictive relevance. All this showed that the ability of deepfake resilience in digital banking had not relied on one technological ability only but instead had been multimodal, real-time, and explanatory modalities—operating as complementary organizational AI abilities. This is in line with the latest financial sector research, highlighting the importance of multimodal deepfake analysis for real-time threat detection, adaptive cybersecurity, and explainable AI as effective countermeasures against AI-powered financial fraud

Both hypotheses were confirmed in a positive correlation between the AI-Driven Multimodal Deepfake Detection Capability (AMMDFC) and the DBFCR ($\beta = 0.268$, $p < 0.001$), suggesting that the integration of all four signals (facial, voice, image and other signals) had enhanced banks' ability to prevent and respond to deep fake frauds. It was also in keeping with Sharma & Sharma, (2024) highlighting the need for multimodal deepfake analysis to the point that it is currently a key component of contemporary deepfake forensics and Gupta et al. (2024) who pointed to multimodal deepfake analysis as one of the pillars of adaptive cybersecurity, particularly in financial services. The same applies to who stressed the need to use multimodal fraud detection using voice biometrics and other signals. The relatively small magnitude of f2 effect of AMMDFC on DBFCR (0.148), however, indicated that an ability to operate multi-modally had not been enough to build strong organizations. This allowed a qualification to studies that, while technically oriented, put a focus on updating the detection architecture as a principle means of combating fraud. It is possible that these differences warrant inclusion, as technical studies would look at only the accuracy with which the events are detected, while organizational resilience would also take into account response processes, human oversight, governance, authentication controls, incident management and operational recovery procedures.

The medium effect size of ARTDDC ($f^2 = 0.235$) supported H2 with a noticeable positive effect on DBFCR ($\beta = 0.344$) ($p < 0.001$). The outcome revealed that being alerted to fraudulent synthetic bank content during transactions had

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

improved the alleviating effect banks had on deepfakes banking fraud. This result agreed with that highlighted on identifying adversarial threat during real-time in the context of Financial Services and Tahraoui et al. (2023) which correlated between Real-time Fraud Detection and secure digital-service environs. Pashentsev & Bazarkina (2023a) also emphasized suspicious-activity detection in real-time as a crucial aspect of the new generation in the realm of biometric cyber security. The outcome also made an important distinction from research mostly focused on the offline forensic detection in the presence of pre-manufactured manipulated content. In digital banking, this lag in detection might undermine the usefulness of security measures, as fraudulent onboarding or authentication processes or transactions may take place quickly. Hence, it is argued that the impact of ARTDDC might be attributed to its potential for stepping beyond the realm of retrospective forensic analysis to a proactive position in the actual risk situation.

The direct effect of XAIDDC on DBFCR was also positive and relatively small ($f^2 = 0.098$) as in DFR1, the effect size is very small, while it reached $\beta = 0.220$, $p < 0.001$ and to supportH3. The finding was corroborated by Das (2024), who highlighted the importance of explainability in bridging the gap between findings in deepfake detection and meaningful interpretation who included explainability in a cyber-fraud detection architecture for use in digital KYC processes. Trustworthy situations of Artificial Intelligence have also been defined as having explainability, transparency, governance and human oversight as conotential factors in the recent study of the financial industry. However, the “relatively small," direct effect did allow for qualification of views that include explainability as a primary security-performance mechanism. Explainability does not exist in the context of proactively identifying or preventing attacks, but instead enables understanding and actionability of AI decisions. The benefits of its resilience may therefore lie in whether these explanations are used up to the potential by the security analysts, fraud investigators and decision makers. The development of AI technologies posed the question of whether AI is going to serve as a science-fiction or reality-based solution to socially nuanced problems, which aligned with the research findings that pointed to ongoing conflicts regarding the sophistication of AI, the degree of explainability, efficiency, and timeliness (Situmeang & Pudjiastuti, 2024).

The impacts on DTDE strengthened evidence on the mechanism by which these capabilities did (or did not) work. AMMDFC significantly influenced DTDE ($\beta = 0.398$, $p < 0.001$; $f^2 = 0.344$), supporting H4. The discovery suggested that using a variety of information modalities had enhanced the strength of the finance industry's identification and detection of synthetic content and/or fake identities. This helped the literature of deepfake forensic techniques being developed at a contemporary stage, which did not rely on a single modality (visual or audio) but multimodal approaches. The result might be due to the complementarity of information across modalities: manipulation which was convincing when only one modality was considered might lead to some inconsistencies when all modalities were taken into account (kinaesthetic, FD, facial movement, voice, identity information, and context). However, the results cannot be taken as implying that multimodal systems were inherently impervious to attacks based on sophisticated systems. New studies have highlighted other potential nuances of reliable detection – including adversarial manipulation, model generalization, new deepfake techniques, and the changing nature of synthetic content – as discussed in recent literature. Hence, in the present findings, multimodality has proved to be effective in the detection, and the efficacy of detection has not been significantly curtailed, though model adjustment has to be done continuously.

ARTDDC had the strongest effect on DTDE ($\beta = 0.426$, $p < 0.001$; $f^2 = 0.395$), supporting H5. According to this finding, one of the critical factors to defend against threats was the bank's pace in identifying the threats as a deepfake, especially in the case of these originating from inquisitive customers.The key factor in the identification of the threat was in how fast banks responsive identified it, particularly in regards to that from inquisitive customers, this result showed. It surmises the real-time adversarial identification and adaptive cyber security in financial services set out, and the real-time identity related detection in a digital KYC risk-based assessment proposed by NIȚĂ1 et al, (2023). It might be appropriate because of the time-sensitive nature of digital banking, such as signing up, biometric authentications, initiating transactions, engaging with customer support, and logging into accounts are all within moments. Thus, too slow response times may not allow for any preventative benefit, even with a technologically correct system. Concurrently, research on the actual applications of AI in real time has noted various challenges, including computation limitations, losses in scalability, privacy concerns, and operational hurdles. The present

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

findings thus suggested that real-time capability was high consequence while existing literature indicated that for implementation, the speed at which it is detected, the accuracy of detection, privacy and operational cost were all important to get right.

XAIDDC had also exerted a strong positive influence on DTDE ($\beta = 0.412$, $p < 0.001$; $f^2 = 0.369$), supporting H6. This indicated that explainability played a role in making threat detection more effective by providing the ability for personnel to comprehend, validate and act on AI alarms. Explainable multimodal machine-learning models for the identification of targeted vishing and social-engineering attacks, and explainability as how to connect between automatic deepfake detection and forensic analysis. This aligns with the broader financial AI literature that calls for explainability in the context of human oversight for critical decision-making processes. But research on explainability has pointed to the introduction of complexity in explanation-provided mechanisms and to the fact that explanatory mechanisms do not necessarily imply trustworthy decision making. This apparent difference can be explained by clarifying technical explainability from organizationally useful explainability; the value of explanations for security depends on whether they can be understood, on the timeliness of their creation, on their accuracy, and on whether they are part of security analysts' decision-making process, or are presented as a technical output.

In support of H7, DTDE was shown to have a significant positive effect on the DBFCR ($\beta = 0.357$, $p < 0.001$; $f^2 = 0.162$), as shown in the Tablet 1 results. This was a key outcome of the proposed framework as it showed that the successful detection of deepfake impersonation, synthetic identities, voice cloning and manipulated content had led to increased fraud and cybersecurity resilience. It resonated with recent investigations that have called for increased financial-system protection by deploying AI in the fight against fraud and identity theft (Srivastava, 2024) and with studies in financial risk management, where AI technology was identified as a component of a financial-system's resilience strategy. However, our findings indicated that threat detection may not be the only factor contributing to resilience as the medium (as opposed to large) effect was suggested. This technology-focused view was given this authority because it proved that it was not necessarily the case that a network would be considered “secure” once it was known an attack had been detected, mitigated and recovered from. Other aspects of resilience also need to be considered such as incident response, access control, employee capability, governance, customer verification, regulatory compliance and recovery. This gave a plausible explanation to why the contributions to DBFCR from DTDE have been substantial up to now, but unfinalised.

This final mechanism was further defined by the findings that came from the mediation. The AMMDFC and ARTDDC partially mediated the relationship between DBFC and CR and XAIDDC and DBFC, respectively (βindirect = 0.142, VAF = 34.63% and βindirect = 0.152, VAF = 30.65% and βindirect = 0.147, VAF = 40.05%), which aided in the support of H8, H9, and H10. The results showed that the three weaknesses of AI capabilities contributed to the overall resilience, likely through an effect on their ability to detect deepfakes which was largely direct in addition to other effects remained unclear. The findings were consistent with the latest significant findings in the literature on PLS-SEM, as significant direct and indirect effects and VAF values of 20%-80% were known to correspond to partial mediation. In substance, the most mediated proportion indicated that there was a relatively greater direct pathway to resilience in the capability of explainability and XAIDD's contribution had towards overall effectiveness of detection, while multimodal and real-time capabilities had relatively stronger direct pathways to resilience. This pattern was used to understand some of the reasons that XAIDDC had had only a minor direct f^2 effect on DBFCR, and a strong effect on DTDE.

The results cumulatively contributed to the existing research in deepfakes and financial-cybersecurity, showcasing the simultaneous use of multimodal, real-time and explainable AI as interchangeable technological features. Rather they had been to express complementary attributes that had been playing different roles in threat success and in the field of total resilience. The findings mostly embraced recent literature in favour of multimodal detection, the timely identification of threats, the provision of explanation oriented guidance for decision making and the adaptive approach to cybersecurity in the financial services sector, but also substantiated more technically oriented views as the direct effects on organizational resilience were small, up to medium, in magnitude. This was significant, however, because many of these partial mediation findings suggested that while they had the potential for sophisticated AI capabilities, this did not necessarily lead to cybersecurity resilience: part of the value had been translated into effective threat detection. Meanwhile, the capacity of banks to withstand the attacks appeared to have actually relied

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

on additional security and organizational factors.Furthermore, the mere fact that mediation was partial suggested that banks' vulnerability had not been just to withstand their attacks, but to a lack of it. The findings therefore affirmed the sociotechnical view of deepfake defense, where the ability of digital banking systems to resist a deepfake attack was determined by both the technology's sophistication and the time needed for deepfakes to be detected, as well as their interpretability, the involvement of the human decision maker, the governance architecture and the resilience of the organization.

THEORETICAL IMPLICATIONS

The research had multiple theoretical implications (a) for research on AI-enabled Digital Banks Cyber Security - (b) for research on Deep Fake Detection - (c) for research on Digital Banking Resilience. First, the results confirmed the Resource-Based View by showing that the ability of AI might be perceived as an organization's different resources, not a generic ability in technology. The potential of AMMDFC, ARTDDC and XAIDDC to significantly impact DBFCR clearly suggests different areas of analytical capabilities that banks could enhance their fraud prevention and cybersecurity capabilities, namely in the capabilities of multimodal analytical breadth, real-time responsiveness and explainability. This infers that multimodal fusion, real-time detection, and adapting AI architectures are vital elements in resilient fraud detection, aligning with the growing body of literature in financial security. Second, the strong effects of ARTDDC ($\beta = 0.426$) and XAIDDC ($\beta = 0.412$) on DTDE added a fresh dimension to the Dynamic Capabilities Theory as it suggested that the added value of the AI resources was not just on the fact of having them, but on their ability to deploy quickly, understand the output of the AI, and refine detection processes accordingly. It mirrored recent writings in the fields of finance and fintech that highlightAI's real-time processing and adaptability as desirable features for financial systems.It reflected recent literature in the fields of finance and fintech, which focuses on adaptability and real-time processing as desired traits of AI-enabled financial systems (Mahammad et al.).

Third, the results built upon Organizational Information Processing Theory by highlighting DTDE as an important process that facilitated the transformation of AI capabilities to cybersecurity outcomes for the organization. DTDE has also made a significant contribution to DBFCR with a VAF value of 35.73 % and partially mediated the effects of AMMDFC, ARTDDC, and XAIDDC on DBFCR, with a 34.63 % VAF, 30.65 % VAF, and 40.05 % VAF, respectively. The findings of this study thus indicated that only gathering and processing richer information was not enough, however; gathering and processing information plus information-processing capabilities capable of increasing the bank's awareness of identity recognition of synthetic identities, manipulated media, voice cloning and deepfakes in the bank created organizational value. This theoretical explanation aligned well with the current studies that focused on integrating fraud mitigation and cybersecurity processes, adaptable detection, and unified multimodal reasoning. Last, the important DTDE–DBFCR relationship brought organizational Resilience Theory to a new level by introducing the effective detection of threat as a source of resilience, instead of considered just a recovery function after the incident. For this reason, it has been necessary to observe and analyze the phenomenon of AI-driven financial cybersecurity as a social-technical process, and this study looked at a selection of such aspects within a single framework instead of viewing these in a purely technical way.

PRACTICAL IMPLICATIONS

The results offered actionable information and insights for banks, cybersecurity managers, fraud-management teams, technology vendors, and risk professionals. Banks should not go by "one size fits all" and should create systems fairly integrated that combine data from facial/video, voice/audio, image, biometric, behavioral, and contextual data. Significant impacts of AMMDFC on DTDE ($\beta = 0.398$) and DBFCR ($\beta = 0.268$) suggested that AMMDFC abilities would help identify threat and general resilience. The recommendation was similar to that of Gilbert & Gong (2024) who demonstrated the potential of multimodal metadata fusion for detection of deepfakes used in a financial multimedia system's synthetic identity fraud, and stressed multimodal authentication responses to synthetic ID threats. The challenge is for banks to use deepfake analytics within the digital onboarding, KYC, biometric authentication, transaction monitoring, and customer-service processes instead of it being an independent security solution.

A strong relationship was seen between ARTDDC and DTDE ($\beta = 0.426$; $f^2 = 0.395$) indicating that the detection infrastructure of banks should be made real time. Deepfake screening should take place during the following stages:

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

at time of onboarding, at time of authentication, on payment authorization, account recover, video banking and high-risk customers' experience, where if the deepfakes are detected, then the suspicious activities can be escalated and losses avoided. The research in recent times also pointed out real time adversarial notification and adaptive cybersecurity measures in financial service domain and real time data integration is crucially important in financial services for resilient cyber security against financial frauds. It is therefore important that a detection engine be integrated with automated alerts, transaction holds, step-up authentication, analyst escalation and incident-response workflows at the bank. But when it comes to automation, it must be adjusted to deal with bogus messages and ensure that real customers aren't affected.

The great influences of XAIDDC on DTDE (β = 0.412) and DBFCR (β = 0.220) revealed that banks should not assess the ability of a deepfake detection system solely based on preserving accuracy of prediction. Security analysts and Fraud Investigators must have understandable evidence of the reasoning why the interaction was flagged as suspicious such as voice/face emotions that don't align, unusual features, manipulated frames etc., etc., all in relation to an identity signal. Explainability might aid in human validation, security auditability, investigation, and defensible security decisions. This used to be in line with the actual guidance given to the financial sector, highlighting good corporate governance, data controls, and integration with cybersecurity, and earlier studies suggesting explainable and adaptive fraud-detection architectures (Tahraoui, Krätzer, & Dittmann, 2023). XAIDDC presented interesting results, with possibly greater demonstration on its direct benefit on DTDE than on its direct influence on DBFCR, which means that explainability should not be simply considered as a compliance tool but as an instrument that will augment the quality and actionability of the threat detection process in banks.

Last but not least, the mediation outcomes have shown that whereas having high-tech tools does not ensure cyber security, another crucial key being acquiring these tools is essential. Banks are investing in hiring qualitative analysts to match their technological capabilities, ensuring you're constantly monitoring models, trying to stress test your system, putting in some tackling incident response integration tasks, providing training for their workers, managing data governance and coordinating with other departments such as the cyber security, fraud, compliance, risk and digital-banking groups. The partial mediation revealed that portions of the effects of the three capabilities of AI on DBFCR were explained by DTDE, which suggested that technological detection may only partially account for the organizations' detection and control of the two other capabilities of AI. This was similar to that AI-technology support for security operations is essential for synthetic fraud; Ghimire (2024) who argued that governance and supervisory controls are important in the context of AI for financial services. The take-home message was for digital banking institutions in the USA to see deepfake defense not as a separate AI application, but rather as an organic function of an organization—including technology for detecting fakes, deploying and intervening in real time, explainability, human oversight, governance, and recovery processes.

CONCLUSION

This research sought to tackle the emerging challenge of deepfake-enabled fraud and cybersecurity risks in U.S. bank digital systems and more specifically a lack of empirical knowledge of how specific organizational AI abilities converted to profitability and how this capability contributed to overall system cyber resilience. While a number of previous studies investigated an increasing number of multifaceted aspects around the deepfake detection, the synthetic identity fraud, multimodal authentication and AI related aspects on financial resilience, the literature still did not offer a more coherent analysis on the role of multimodal, real-time and explainable Artificial Intelligence within the organizational framework and how such capabilities contributed to financial resilience. As digitization and AI tools, such as digital identities, voice cloning, manipulated videos, and more were novel vulnerabilities in digital onboarding, authentication, transactions, and customer interactions, this gap became more significant.

An outcome of the results is found to be a strong support of the proposed framework of the study. For DBFCR, AMMDFC (β = 0.268), ARTDDC (β = 0.344) and XAIDDC (β = 0.220) have each had a significant positive effect. DTDE also increased significantly, with a β of 0.398 for the capability AMMDFC, a β of 0.426 for capability of ARTDDC and a β of 0.412 for the capability of XAIDDC. A higher DTDE also had a significant positive impact on DBFCR with a β value of 0.357. So, we have found evidence for all seven direct hypotheses. The model had explained 54.2% of the variance in DTDE and 64.0% of the variance in DBFCR which showed that there was meaningful

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

explanatory power of the models. These results provided support for the recent arguments for require of multi-modal information processing, quick and adaptive threat detection, and interpretable AI for effective financial cybersecurity; instead of only relying on own detection mechanisms.

The results of the mediation were used to further explain the contribution of these capabilities to resilience. It showed that DTDE explained part of the effect of AMMDFC, ARTDDC and XAIDDC on DBFCR; the effects of AMMDFC, ARTDDC and XAIDDC were 0.142, 0.152 and 0.147, respectively and the values of VAF were 34.63%, 30.65% and 40.05%, respectively. However, the results revealed that AI did not directly increase cybersecurity resilience – but rather that building effective cybersecurity capabilities to detect threats tied to deepfakes was essential to the contribution AI was making to strengthening cybersecurity capabilities. Simultaneously, the clues from partial rather than full mediation indicated that other banking resilience factors like governance, authentication controls, incident response, human expertise, organizational coordination and recovery capabilities were also likely to be important to banking resilience. This interpretation was in alignment with the current investigation highlighting the necessity of incorporating AI architectural integrations, real-time data, organization controls, regulatory and governance protocols in determining resilient financial security.

This theoretical integration of the RBV, DC, OIP and OR focused on how the differentiated AI resources can be converted to AI resilience outcomes by proper threat detection. Notably, the study provided position-specific explanations where technological capabilities are linked with organizations' cybersecurity capabilities, moving the research agenda beyond the level of algorithm detection accuracy provided to the level of organizational capabilities. In practice, the results identified that banks need to build a holistic architecture for defending against deepfakes; one that incorporates multimodal analytics, real-time detection, explainable outputs, human supervision, continuous surveillance and incident response. The impact that ARTDDC and XAIDDC had particularly had on DTDE further highlighted the need for detection systems to be technically capable, but also quick and easy to interpret in order to facilitate decision making for operational security. These implications aligned with the growing body of financial-security research that emphasized the need to adopt real–time, multimodal, adaptive, and explainable means of fraud defense utilizing AI techniques.

The study overall shows the need to see increasing one's resilience to DFF not as a result of simply implementing new and sophisticated financial fraud detection algorithms, but as a sociotechnical capacity. Multimodal detection broadened threat information banks have, real-time detection allowed them to respond in time, explainable AI enhanced the interpretability and actionability of the decisions they made with detection, and DTDE enhanced use of these capabilities into more effective fraud prevention and cybersecurity results for banks. The findings suggested that for U.S. digital banking institutions to be more capable of detecting, resisting, reacting to, and recovering from new threats created by deepfakes, they must adopt the technology's sophistication and sophistication processes, judiciously supplemented by technology-driven responses, governance, and organizational reaction protocols. This holistic view was the main finding of the research and served as a basis for additional empirical studies to test the capabilities of AI in financial fraud deterrent systems and the resilience of financial systems against cybercrime in a more synthetic and automated financial landscape.

LIMITATIONS

Although the study contributed, there are some limitations that should be taken into consideration whilst interpreting the results of this study. In the first regard the research design was a quantitative cross sectional, which examined the relationship between AMMDFC, ARTDDC, XAIDDC, DTDE, and DBFCR at one time. While the proposed relationships were supported by the PLS-SEM results, experimental design did not identify temporal causality or identify how the ability of banks to detect synthetic media and their cybersecurity capability changed over time with the evolution of the synthetic media technologies. Secondly, perceptual data in this study was obtained from a structured questionnaire, which involved self-reports. Measurement and common-method assessments indicated that there were no significant common-method concerns with these reported assessments, though respondents' responses may have been influenced by their knowledge, experience, recall or subjective evaluation of their organizations' cybersecurity capabilities. Third, the study examined digital banking transactions in the United States, and therefore these results may not apply directly to financial systems in other countries with varying technological

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

and cyber security systems, financial regulatory frameworks, and institutional environments. Fourth, the model focused primarily on three capacities of AI; Multimodal, real-time, and explainable deep-fake detection were not included in the model since they may be relevant for separate analyses.Fourth, only three of the abilities of AI, including Multimodal, real-time, and explainable deepfake detection are within the scope of the model, while other potentially relevant aspects were not covered (such as adversarial robustness, threat intelligence, employee cybersecurity competence, AI governance, data quality, privacy controls, and regulatory readiness). This limitation was crucial because modern literature has highlighted the need for resilience in countering deepfakes, which requires adapting to the growing complexity of fakes, governing, authenticating and integrated response. Last, the study didn't directly compare the detection effectiveness and resilience of an organization against each of the various deep fake detection algorithms, but rather its perceptions; therefore, the results would serve as an indicator into organizational AI capability, but not necessarily actual detection success or failures against actual attack datasets.

FUTURE RESEARCH DIRECTIONS

These limitations can be tackled in future research through using longitudinal designs to investigate changes in deepfake detection over time, and to see if the three scores, DTDE, ARTDDC and XAIDDC, change consistently with improvements in the detection of deepfakes. Survey data could be supplemented by the findings of experimental and quasi-experimental research, which would enable the actual success rates, response times, false positive rates and incident impacts at which professionals or security systems in banks could be evaluated through carefully controlled setups based on a set of representative, concrete scenarios involving fauxbrevo, imitations of well-known brands or familiar voices. Comparative analyses within and between banks, fintech companies, credit unions and other financial institutions might also be possible – as well as expanding the model to other countries to examine if the identified relationships are moderated by the regulatory, technological and cultural conditions. The latter is especially apt as recent financial-security research has focused on the relationship between technological innovation, risks of fraud, and regulatory reactions.

The conceptual framework can be further expanded and deepened through further studies that include AI governance, human–AI collaboration, culture of cybersecurity, capability of threat-intelligence, adversarial robustness, protection of privacy, transmission of expertise from employees, and readiness of regulation as moderators, mediators or antecedents. A major aspect to concentrate on is adversarially robust detection and continuous adaptation to new attacks, given the potential that the fast-evolving deep fake technology can jeopardize the generalizability of the detection system to attacks that it has never encountered. A multimodal, real-time, explainable AI can be done alone or in combination; researchers could explore whether there are synergy effects and complementary effects between the modalities of the collection than only examine their singular, individual contributions. Furthermore, future research might examine the detection value of individual modalities (facial video, cloned voice, simulated identity documents, behavioural biometrics, and inconsistencies among modalities) to establish their detection ranges for various processes within the banking sector. Many related studies related to the multimodal deepfake fraud detection and federated deepfake fraud detection can be useful for this investigation (Abdeltawab, 2024).

Finally, future studies would benefit from investigating the trade-off between effectiveness of detection and operational factors of computational cost, operational friction, privacy, explainability, high false alarms and response time. The qualitative interviews and mixed method studies could also disclose certain organizational processes that quantitative studies might not capture as fully.Qualitative interviews and mixed method studies involving fraud investigators, cybersecurity analysts, compliance professionals, AI developers, and banking executives could identify components of organization processes that are not well represented in quantitative studies. Additional objective organizational metrics, such as recorded incidents of deepfakes, successful mitigation of fraudulent transactions, financial losses, detection delay, false positives, recovery time and the efficiency of time spent reacting to an incident, might also be added in future research. Linking perceptual measures with operational cybersecurity information would deliver more robust information on the real-world implications of AI deepfake detection for measurable cyber-resilience. With the ongoing development of deepfake technologies, research needs to additionally focus on extending the validity of AI to help detect synthetic frauds to a longitudinal, multimethod, cross-national, and real-world

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

approach, not just on the issue of its ability to detect those synthetic frauds, but also the ability to remain adaptive, interpretable, trustworthy, and operationally effective in an ever changing digital banking landscape.

REFERENCES

[1] Abdeltawab, M. S. (2024). Ethical Challenges and Mitigation Strategies for Deepfakes: The Role of Explainable AI and Legal Frameworks. International Conference On Artificial Intelligence, Computing, IOT and Data Analytics,

[2] Akpabio, E., & Narad, S. (2024). Artificial Intelligence and Cybersecurity: Challenges, Opportunities, and Defensive Techniques. International Conference on ICT for Sustainable Development,

[3] Aleksandrova, A., Ninova, V., & Zhelev, Z. (2023). A survey on ai implementation in finance,(cyber) insurance and financial controlling. Risks, 11(5), 91.

[4] Ali, A., Khan, M. A., Farid, K., Akbar, S. S., Ilyas, A., Ghazal, T. M., & Al Hamadi, H. (2023). The effect of artificial intelligence on cybersecurity. 2023 International Conference on Business Analytics for Technology and Security (ICBATS),

[5] Ali, A., Septyanto, A. W., Chaudhary, I., Al Hamadi, H., Alzoubi, H. M., & Khan, Z. F. (2022). Applied artificial intelligence as event horizon of cyber security. 2022 International conference on business analytics for technology and security (ICBATS),

[6] Anagha, R., Arya, A., Narayan, V. H., Abhishek, S., & Anjali, T. (2023). Audio deepfake detection using deep learning. 2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART),

[7] Bhardwaj, A., & Goundar, S. (2024). Rise of the Shadows: Profiling the New Age Cybercriminals. In Corporate Cybersecurity in the Aviation, Tourism, and Hospitality Sector (pp. 18-39). IGI Global Scientific Publishing.

[8] Bheemisetty, N. (2024). From fragmentation to agility: Nautilus architecture for risk management modernization. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 7(4), 10673-10682.

[9] Brooks, C. (2024). Inside cyber: how AI, 5G, IoT, and quantum computing will transform privacy and our security. John Wiley & Sons.

[10] Capuano, N., Fenza, G., Loia, V., & Stanzione, C. (2022). Explainable artificial intelligence in cybersecurity: A survey. IEEE Access, 10, 93575-93600.

[11] Das, R. (2024). Generative AI: phishing and cybersecurity metrics. Crc Press.

[12] Dewani, N. D., Khan, Z. A., Agarwal, A., Sharma, M., & Khan, S. A. (2022). Handbook of research on cyber law, data protection, and privacy. IGI Global.

[13] Engvall, N. (2024). The Influence of Institutional Factors on AI adoption in EU banking cybersecurity:: A narrative literature review.

[14] [Record #1480 is using a reference type undefined in this output style.]

[15] Folds, C. L. (2022). How Hackers and Malicious Actors Are Using Artificial Intelligence to Commit Cybercrimes in the Banking Industry. Colorado Technical University.

[16] Geradts, Z., & Franke, K. (2023). Artificial intelligence (AI) in forensic sciences. John Wiley & Sons.

[17] Ghimire, A. (2024). Harnessing big data with ai-driven bi systems for real-time fraud detection in the us banking sector. BULLET: Jurnal Multidisiplin Ilmu, 3(6), 731-743.

[18] Gilbert, A., & Gong, Z. (2024). Digital identity theft using deepfakes. In Information technology security and risk management (pp. 307-314). CRC Press.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[19] Gupta, S., Goyal, N. K., & Kumar, A. (2024). A Historical, Present, and Prospective Review of Artificial Intelligence's Role in Securing Personal Information and Private Data. Artificial Intelligence and Communication Techniques in Industry 5.0, 257-272.

[20] Hämäläinen, M. (2024). Analysis of artificial intelligence in cybersecurity identity and access management: potential for disruptive innovation.

[21] Horowitz, M. C., Allen, G. C., Saravalle, E., Cho, A., Frederick, K., & Scharre, P. (2022). Artificial intelligence and international security. Center for a New American Security.

[22] Kaushal, A., Kumar, S., & Kumar, R. (2024). A review on deepfake generation and detection: bibliometric analysis. Multimedia Tools and Applications, 83(40), 87579-87619.

[23] Kaushik, K., Khan, A., Kumari, A., Sharma, I., & Dubey, R. (2024). Ethical considerations in AI-based cybersecurity. In Next-generation cybersecurity: AI, ML, and Blockchain (pp. 437-470). Springer.

[24] Kazimierczak, M., Habib, N., Chan, J. H., & Thanapattheerakul, T. (2024). Impact of AI on the cyber kill chain: A systematic review. Heliyon, 10(24).

[25] Kolade, T. M., Aideyan, N. T., Oyekunle, S. M., Ogungbemi, O. S., Dapo-Oyewole, D. L., & Olaniyi, O. O. (2024). Artificial intelligence and information governance: Strengthening global security, through compliance frameworks, and data security. Available at SSRN 5044032.

[26] Kovací, P.-D. (2023). Threat actors seeking to exploit ai capabilities. Types and their goals. Strategic Impact, 89(4), 53-63.

[27] LOZONSCHI, C., & BAKHAYA, I. (2023). Artificial Intelligence and its Impact on Cybercrime. International Conference on Cybersecurity and Cybercrime,

[28] Ma, K. W. F., Dhot, T., & Raza, M. (2023). Considerations for using artificial intelligence to manage authorized push payment (APP) scams. IEEE engineering management review, 51(3), 166-179.

[29] Mahammad, F. S., Geetha, G., Thanusha, G., Manasa, A., Harika, D., & Jahnavi, K. Deepfake Detection Using Hybrid Models. Proceedings Copyright, 911, 922.

[30] Mendoza, C. M. C. (2024). Fusing Generative AI and Agentic Workflows for Fraud Mitigation in InsurTech. The American Online Journal of Science and Engineering (AOJSE), 2(02).

[31] Mienie, E. L., & Payne, B. R. Developing a Disinformation Incident Response Playbook: Combatting Real-Time Disruption via Deepfakes and Generative AI.

[32] Nibigira, N., Havyarimana, V., & Xiao, Z. (2024). Artificial intelligence adoption for cybersecurity in Africa. Journal of Information Security, 15(2), 134-147.

[33] NIȚĂ1, G., GĂBUDEANU18, L., CERNAEANU19, C. C., RAICU20, G. M., & ȘCHEAU21, M. C. (2023). Artificial Intelligence and Key Risk Indicators in Cyber Frauds Prevention. Journal of Financial and Monetary Economics, Centre of Financial and Monetary Research "Victor Slavescu, 11, 92-111.

[34] Nnamdi, N., Oniyinde, O., & Abegunde, B. (2023). An appraisal of the implications of deep fakes: The need for urgent international legislations. American Journal of Leadership and Governance, 8(1), 43-70.

[35] Pashentsev, E., & Bazarkina, D. (2023a). Malicious use of artificial intelligence and threats to psychological security in Latin America: common problems, current practice and prospects. In The Palgrave handbook of malicious use of AI and psychological security (pp. 531-560). Springer.

[36] Pashentsev, E., & Bazarkina, D. (2023b). Malicious use of artificial intelligence: Risks to psychological security in BRICS countries. In The Palgrave handbook of malicious use of AI and psychological security (pp. 297-334). Springer.

[37] PASHENTSEV, E., BAZARKINA, D., MIKHALEVICH, E., & WONG, N. (2024). The Malicious Use of AI: Challenges to Psychological Security in the People's Republic of China4. ББК 16.6 М21, 62.

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

[38] Pavel, T. (2023). Regulating AI-based Cyber Attacks: Exploring multi-dimensional approaches and challenges. Artificial Intelligence: Crime, War, and Justice, 217.

[39] Pleshakova, E., Osipov, A., Gataullin, S., Gataullin, T., & Vasilakos, A. (2024). Next gen cybersecurity paradigm towards artificial general intelligence: Russian market challenges and future global technological trends. Journal of Computer Virology and Hacking Techniques, 20(3), 429-440.

[40] Rafy, M. F. (2024). Artificial intelligence in cyber security. Available at SSRN 4687831.

[41] Reddy, A. S. (2023). Future Proofing Insurance Operations: A Guidewire-Centric Approach to Cloud, Cybersecurity, and Generative AI.

[42] Saalbach, K. (2023). Artificial Intelligence and Cyber Attacks.

[43] Sailaja, P. (2023). Legal And Ethical Considerations of Managing Identities, Ranging from Aadhaar to Global Digital Ids. Anusandhanvallari, 1165-1173.

[44] Sarcea, O.-A. (2024). AI & Cybersecurity–connection, impacts, way ahead. International Conference on Machine Intelligence & Security for Smart Cities (TRUST) Proceedings,

[45] Schiliro, F. (2024). From Crime to Hypercrime: Evolving Threats and Law Enforcement's New Mandate in the AI Age. arXiv preprint arXiv:2411.10995.

[46] Schreiber, A., & Schreiber, I. (2024). Bridging knowledge gap: the contribution of employees' awareness of AI cyber risks comprehensive program to reducing emerging AI digital threats. Information & Computer Security, 32(5), 613-635.

[47] Sharma, A., & Sharma, R. (2024). Generative artificial intelligence and legal frameworks: Identifying challenges and proposing regulatory reforms. Kutafin law review, 11(3), 415-451.

[48] Shrivastava, G., Ojha, R. P., Awasthi, S., Bansal, H., & Sharma, K. (2024). Emerging threats and countermeasures in cybersecurity. John Wiley & Sons.

[49] Situmeang, S. M. T., & Pudjiastuti, D. (2024). Adequacy of Indonesian Cybercriminal Law Against Generative AI-Based Social Engineering Attacks. Res Nullius Law Journal, 6(1), 82-97.

[50] Srivastava, D. A. K. (2024). Artificial Intelligence, Technology and Cyber Law: Challenges and Future Prospects. Integrity Education India.

[51] Staynings, R. (2023). Cybersecurity. In Digital Health Entrepreneurship (pp. 131-155). Springer.

[52] Tahraoui, M., Krätzer, C., & Dittmann, J. (2023). Defending informational sovereignty by detecting deepfakes: Risks and opportunities of an AI-based detector for deepfake-based disinformation and illegal activities. Weizenbaum Conference" Practicing Sovereignty: Interventions for Open Digital Futures",

[53] Tahraoui, M., Krätzer, C., Dittmann, J., & Aden, H. (2023). Defending Informational Sovereignty by Detecting Deepfakes? Opportunities and Risks of an AI-Based Detector for Deepfakes-Based Disinformation and Illegal Activities. Weizenbaum Journal of the Digital Society, 3(2).

[54] Udeze, C. T., Falebita, O., & Fasasi, S. T. (2024). AI-Based Cybersecurity Risk Prediction and Management. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 16(04), 216-225.

[55] Waizel, G. (2024). Bridging the AI divide: The evolving arms race between AI-driven cyber attacks and AI-powered cybersecurity defenses. International conference on machine intelligence & security for smart cities (TRUST) proceedings,

Copyright © 2024 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.