INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs.2026180207
Received: 15.02.2026 | **Accepted:** 11.06.2026 | **Published Online:** 16.06.2026

-RESEARCH ARTICLE-

THE ROLE OF INTERNAL AUDITING IN AI RISK MANAGEMENT FOR DEEPFAKE EVALUATION: A STUDY OF COSO AND AIRMF FRAMEWORKS

Tiba Abdul Karem Mohamed Jafar
ORCID: https://orcid.org/0009-0003-0734-835X
Email: tibaabdulkareem@uomustansiriyah.edu.iq

Ibithaj Ismail Yaqoob
ORCID: https://orcid.org/0000-0001-8543-4112
Email: hussainalaa10000@uomustansiriyah.edu.iq

Zahra Hasan Oleiwi
Al-Mustansiriyah University — College of Administration and Economics — Department of Accounting
ORCID: https://orcid.org/0000-0003-3850-731X
Email: zahra_alamiri65@uomustansiriyah.edu.iq

Jaafar Abdulhussein Hiloalkiabi
Ibn Sina University for Medical and Pharmaceutical Sciences, Iraq.
ORCID: https://orcid.org/0009-0003-5571-4617
Email: Jaafar.abdulhussien@ibnsina.edu.iq

Athmar Abdulrahman Sharhan
Department of Financial and Banking Sciences Al-Bayan University
ORCID: https://orcid.org/0009-0007-4525-5866
Email: Athmar.a@albayan.edu.iq

—Abstract—

This study investigates the factors influencing deepfake risk evaluation in Iraqi banking sectors through the Artificial Intelligence Risk Management Framework (AIRMF), Artificial Intelligence Risk Management (AIRM) practices, Committee of Sponsoring

Citation (APA): Jafar, T. A. K. M., Yaqoob, I. I., Oleiwi, Z. H., Hiloalkiabi, J. A., sharhan, A. A. (2026). The Role of Internal Auditing in AI Risk Management for Deepfake Evaluation: A Study of COSO and AIRMF Frameworks. *International Journal of Economics and Finance Studies*, 18(02), 162-180. doi: 10.34109/ijefs.2026180207

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

Organizations framework (COSO) framework for Internal Auditing (INA). Data from 313 respondents of multiple banking institutions and assessed it using the Smart Partial Least Squares (SmartPLS) two-stage analytical procedure. The empirical results revealed that AIRM and AIRMF generally had statistically significant positive effects on the deepfake risk assessment, whereas INA had a weaker but still positive effect. On the other hand, CFR was not found to be statistically significantly associated with deepfake risk assessment. Based on these findings, the study recommends that Iraqi banking institutions strengthen risk governance mechanisms in favour of artificial intelligence (AI) to provide greater identification and prevention of AI-driven threats such as risks attributable to deepfakes alongside mitigation efforts. The results also show the importance of improving INA procedures to provide effective monitoring, strengthen internal control systems, and ensure compliance with regulations. Further, the incorporation of AI-oriented risk dimensions into any modernisation and adaptation of CFR is envisaged as a key means through which to respond to the changing technological threat landscape. It also highlights the significant potential of sector-wide collaboration, especially in areas such as sharing knowledge, building shared capacity and promoting effective new ways of working to address AI-related risks within sectors. When implemented in tandem, these measures should improve the institutional resilience of establishments against AI-related security threats and enhance the deepfake risk assessment and mitigation mechanisms for the banking industry as a whole.

Keywords: Internal Auditing, AI Risk Management, Deep Fake, AIRMF, COSO.

INTRODUCTION AND LITERATURE BACKGROUND

AI-powered technologies are growing quickly, presenting never-before-seen challenges such as identity manipulation, fraud and online deception (Adavelli et al., 2025; Arif et al., 2024). An entity could unwittingly receive a formal letter from some bank saying they had failed to repay a loan it never took out, or for example is not allowed to travel abroad because official records now falsely state that they have used forged travel documentation Artificial Intelligence (Vecchietti et al., 2025). This may often be the case when cybercriminals use deepfake technology to forge passports or other identity documents with personal information of an innocent person without their knowledge, allowing them to impersonate someone else's identity (Mandal, 2025; Vecchietti et al., 2025). These events over so-called speculative and futures are no longer hypothetical. Recent evidence indicates that deepfake technologies are acquiring a closer relation to serious crime (e.g., human trafficking, extortion, revenge pornography and ransomware activities as well as organized cybercrime) (De Bock et al., 2024; Filieri et al., 2022; Sai & Wang, 2026). As a result, the world is moving into an era of technological disruption (Portuguez-Castro, 2025) in which almost anyone or organization can

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

become a target for AI-based identity deception and synthetic media assault (De Bock et al., 2024).

As technology becomes more advanced, the threats it creates grow larger and more complex (Kavanagh, 2022). The deepfake generation toolset has become more accessible and readily available (Masood et al., 2023), allowing even those with limited technical expertise to produce highly convincing fraudulent/audio/ideological material. The increasing public access to AI-based manipulation technologies will greatly increase the risk of future cyber campaigns (Bharati, 2024) that deceive people who are easy to persuade, as well as create new forms of fraud designed to gain money (Gambín et al., 2024). However, modern deepfakes have become such high quality that even technical experts can often struggle to tell fake from real content (Das et al., 2024). This has both personal and organisational implications as synthetic media is able to manipulate reality by recreating fake events that may appear real in context and naturally. For instance, in one recent case of cyber criminals using this technology a social engineering fraud used deep fake to impersonate a corporate Chief Financial Officer persuading an employee working in the finance department to transfer USD 25 million (Chen, 2024). Likewise, Hong Kong authorities said that criminals used information from stolen identities to create fake loan applications and banking accounts which resulted in the same identity being implicated in 90 loan and 54 bank account applications within a three-month period (Chen & Magramo, 2024; Chen, 2024).

As these deepfake-related threats present themselves, executives and decision-makers in organisational environments are faced with complex strategic dilemmas relating to how they manage the problem of deepfakes. For organisations, the choices may revolve around whether to intervene financially (to mitigate getting involved in attempts to suppress manipulated material) or take the risk of severe reputational damage (Kietzmann et al., 2020). A notorious example occurred in June 2019, when a manipulated video of Mark Zuckerberg emerged on social media showing the Facebook chief discussing his power over massive amounts of personal data (Chen, 2024; Haris, 2020). Completely fake, the altered video looked very real and was meant to confuse public perception of the practices of Meta Platforms on data. Given the rapid advancement and spread of deepfake technologies and the increasing need for advanced detection capabilities and governance system (Vecchietti et al., 2025).

Deepfakes can be very roughly defined as media files (often videos, images or even sound recordings) (Farid, 2022) that have been AI-rendered through complex computer techniques which represent something very realistic-looking and sounding but which is not real (Jayashre & Amsaprabhaa, 2024). These technologies can change many features of a human being their face, age, gender identity, emotion on the facial expression, hairstyle and skin complexion. Prior art usually identifies four main types of deepfakes: face swap (Walczyna & Piotrowski, 2023), facial re-enacting (Nawaz et al., 2024), attribute manipulation (Rehaan et al., 2024) and custom or synthetic

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

generation of the full face (Rehaan et al., 2024). From a historical standpoint, early instances of deepfakes can be dated back to 2017 when it first surfaced through a Reddit user (Güera & Delp, 2018). Since then, deepfake technologies have been extensively employed mostly in attacks against celebrities (Tariq et al., 2022), political figures (Islam et al., 2024), and prominent public personalities (Perot & Mostert, 2020). Take media where Alia Bhatt was manipulated, there are claims that show the actress being part of content that she never made to be used widely on internet (Rajput & Arora, 2023). The video, which was fake, received more than 17 million views on Instagram before being removed (Usmani et al., 2026).

Another escalating form of AI-enhanced manipulation is found in audio deepfakes, where an individual declares statements that the target has never uttered, commonly known as voice cloning or voice impersonation. One infamous instance came during the United Kingdom's Labour Party conference in 2023 when a fake audio of Keir Starmer emerged, a fabricated recording of sounds that made it seem as if the politician had been abusing party staff (Usmani et al., 2026). A similar incident occurred during the 2024 New Hampshire Democratic primary, when over 20,000 voters were said to have received phone calls, generated by AI systems that replicated Joe Biden's voice and told them not to vote (Usmani et al., 2026).

Considerable academic attention has been dedicated to the increasingly relevant role of INA in reducing organisational threats and enhancing governance mechanisms. In previous studies, for instance (Chang et al., 2014), it has been argued that managing risk effectively might include transferring some risks to external partners and assurance providers including auditors. In a similar vein, the Organisation for Economic Co-operation and Development (OECD, 2014) have acknowledged that modern governance regulations and institutional frameworks are increasingly bringing INA to bear on risk governance systems (Coetzee, 2016). Within this scenario, (Coetzee, 2016) explored perceptions regarding the contingently perceived contributions of Chief Audit Executives, audit committee leaders and senior managers about strategies for positioning INA in public-sector risk management environments in South Africa. The research also examined the relationships of INA with other organisational risk management structures and alleged a varying level of collaboration among these governance arrangements. The results showed that Chief Audit Executives held different views than senior management and audit committee chairs. In addition, while formal risk management structures existed, they had little Impact on Stakeholders perception of the adequacy of INA contributions to organisational risk governance.

In addition, Sharma and Selwal (2026) argued that AI technologies are an essential enabler of the creation and distribution of deepfakes using innovative machine learning techniques. These systems can produce far more realistic synthetic media such as video,

audio and multimodal. Generative models for rendered data generation with deepfakes technology advancing at an accelerating rate, and the digital media ecosystem itself expanding exponentially in its complexity, new scholarship calls for a sophisticated analysis that considers both the potential benefits such technologies offer their many dire consequences as well. Throughout history, media manipulation was by no means new; AI and deep fakes have just made it worlds more advanced and mass available. Thus, Sharma and Selwal (2026) performed a survey of the evolution/development deepfake generations & detections from 2017-2025 that entails the latest methods to generate/fake misleading media as well as recognize its forgeries. Modern architectures were studied including adversarial networks, Vision Transformer (ViTs) and hybrid learning models analysing their strengths and weaknesses. It also recognized flaws in the current detection systems, particularly regarding the adversarial attacks or deceptive actions that can easily bypass such detection methods. Two panels also generated insightful dialogue on place of contemporary concepts of AI concerning digital forensic methodologies that were earlier employed in both the validation and authentication of media. In addition to technical elements, the research also looked at broader operational problems like scalability limits and disclosure mandates as part of the social-legal effects study into widespread use of synthetic media. Additionally, this study included discussions about evaluation benchmarks, technology infrastructures and performance metrics, as well as policy recommendations to mitigate the risks of deepfakes. In conclusion, the research highlighted the need for development of more interpretable, adaptive and policy driven detection systems that are ready to respond in a timely manner with AI-generated threats that would be changing rapidly.

The connection of INA and risk governance has also received much academic and practical attention from multiple industrial sectors. For instance, Wang and Li (2011) contended that many major engineering enterprises are getting their internal audit departments to assist in risk governance activities to prevent losses due to both internal and external uncertainties while at the same time increasing enterprise value creation. The research aims to provide insights on not only how engineering organisations can or may adopt INA but also risk-based approaches (e.g., Analytic Hierarchy Process) for risk identification and evaluation that a typical engineering organisation is likely to employ within their risk management processes. The study also evaluated integrated benefits of INA as part of larger risk governance systems and investigated current practices regarding engineering risk management. The findings suggest this necessitates the introduction of more risk-led approaches to auditing for both improved corporate competitiveness and sustained development on the part of economic organisations facing uncertainty in their respective business environments.

In contrast, contemporary research has studied Enterprise Risk Management (ERM) frameworks such as the Committee of Sponsoring Organizations of the Treadway Commission (COSO) and International Organization for Standardization (ISO) systems in a number if ways; they have used a plethora of explanatory and mediating variables

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

affecting outcomes pertaining to governance. For examination, Karanja (2017) explored if firms aligned their ERM existed with the main domains embedded in the COSO 2004 and ISO 31000:2009 frameworks. They focus mostly on alignment with strategy, operational efficiency, reporting reliability and regulatory compliance and maturity of implementation. Findings on ERM found that most organisations have ERM at entity-level or higher and realise this is a strategic resource enabling enhanced operational performance capability as well as enhanced compliance capacity.

This adds to the body of evidence around which INA can deliver in evaluating and analysing risk governance mechanisms for the systems on AI embedded tools emerging from deepfake technologies. An improved comprehension of the implications in the AIRMF and CFR frameworks is therefore crucial to audit systems that not only assess or monitor but also compensate risks associated with AI-based manipulation technologies. It is further argued that the resilience of AI from a governance and oversight angle might require auditing approaches different from traditional ones that can evolve with constantly developing threats. It does so by submitting practical recommendations aimed at improving the risk governance mechanisms for organizations working in a landscape that is increasingly exposed to related risks from deepfakes. As such, the research supports more systematic incorporation of INA into strategic communications strategy targeting new threats from emerging AI-enabled technologies.

Although AI technologies are being deployed extremely rapidly across nearly all sectors of business and in financial institutions worldwide, fears about the risks of deploying AI have continued to mount pressure to make these deployments safer. These developments necessitate frameworks that can help handle unknowns and variances in technology adoption. Internationally, various governance structures have evolved, AIRMF and the new updated COSO framework considered in the present study. On the contrary, although banks in the Iraqi environment increasingly rely on AI-based technologies and naturally face a wide range of threats (including deepfakes), formalized frameworks still be underdeveloped. According to these issues, the current study investigates the following research objectives:

1. Assessed the importance of INA and how it relates to other data available when evaluating technologies in AIRM strategies associated with deepfake.
2. Evaluate the effectiveness of CFR in managing AI system risks related to deepfake.
3. To investigate how AIRMF contributes to more robust mitigation strategies for deepfake oriented AI risks
4. To delve into the mutual influence of AIRM approaches on the integration and execution of CFR principles and AIRMF in assessing incursion threats posed by deepfake technology.

Figure 1 Conceptual model underlying the study illustrating the relationships among INA, AIRM, CFR and AIRMF with deepfake risk assessment.

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

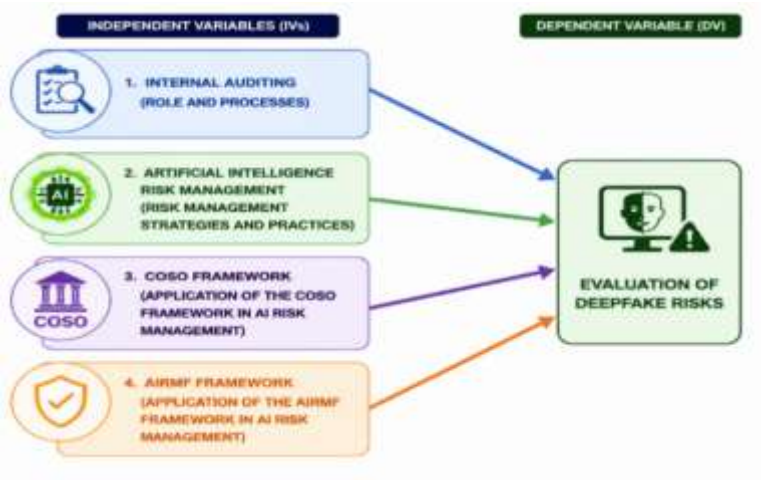

Figure 1: Framework of the Study

RESEARCH METHODOLOGY

This study instead employs a deductive research approach, as it aims to explore the relationships among variables through an instrument-based quantitative exploration. Based on this methodological approach, a well-planned survey tool was conducted by the authors to cover the independent, dependent and demographical variables associated with objectives of the study. The questionnaire primarily focuses on key constructs pertaining to INA, AIRM, CFR and AIRMF with specific focus on the extent of their efforts in governing and quantifying deepfake risks in the banking sector. The questionnaire consisted of several parts. The first part gathered respondents' demographic data, and the remaining parts included measurement items rated using a five-point Likert scale from Strongly Disagree to Strongly Agree. In the section regarding AI applications in INA, respondents were asked to rate statements that indicate if AI-based auditing tools had enhanced the accuracy and effectiveness of risk assessment procedures at their banking institutions.

At the same time, that this was happening, the AIRM section included assertions regarding how processes ought to be set up within banking organizations and what should constitute governance structures and dissemination modes in order to constantly amplify AI-oriented strategies for risk governance around emergent technological risks like supply chain vulnerabilities and other forms of attack surface from deepfake-related attacks. Owing to the structure of this questionnaire, a full analysis of AI adoption with regards to all aspects of auditing and how governance frameworks are embedded within organisations to identify potential risks that arise from deepfakes either directly or indirectly could be undertaken. Various demographic variables are also measured in the study to better understand survey respondents' characteristics. Independent variables were gender, age category, banking position and years of experience. The sample of the study was employees working in private banking mainly those working within INA

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

departments. Thus, the sample was composed of individuals who were directly involved in audit and risk management functions within private banking firms.

After the construction of the questionnaire, a pilot testing process was performed to examine the reliability and appropriateness of the measuring instrument for full-scale data collection. The pilot study utilized a sample consisting of 30 respondents working at the banking industry. The statistical results of the pilot assessment are listed in Table 1 and confirm that all measurement constructs exhibit acceptable reliability. In particular, the Cronbach's alpha coefficients for all items in the questionnaire were higher than their limit acceptable threshold value (i.e. > 0.70), thus confirming internal consistency and reliability of the instrument used in this study.

Table 1: Reliability of the Questionnaire

Nature of the Variable	Name of the Variable	Cronbach Alpha Values
Independent	Internal Auditing (IV)	0.841
Independent	AI Risk Management (IV)	0.812
Independent	COSO Framework (IV)	0.751
Independent	AIRMF Framework (IV)	0.886
Dependent	Evaluation of Deepfake Risks (DV)	0.729

During the initial stage of data collection, researchers used a convenience sampling method focusing on banks that are active in Iraq. We complete the whole process of data collection in a timespan of three weeks. At this stage a grand total of 341 questionnaires were collected from respondents. A screening and validation process identified 28 questionnaires as incomplete or incorrectly filled, preventing further statistical analysis. Thus, these answers were removed in the final dataset. After data-cleaning, the study confirmed a total of 313 valid questionnaires forming the final sample used for empirical analysis.

Statistical Methods Descriptive and inferential statistical analyses were conducted to test the associations between the study constructs as part of the analytical procedures in this investigation. Furthermore, the demographic variables of respondents were analysed by frequency distribution analysis to gain an overall description of the sample. Table 2 provides the demographic outcomes. Description of demographic results showed that the highest percentage of respondents were from age range 41–50 years (128 people). Next 53 respondents in the range of 31–40 years-old followed by 48 respondents under the age of 30. In addition, 45 respondents were in the age category of 51–60 years while 39 participants were aged above the 60 years. Regarding gender distribution, most female respondents corresponded to 59 participants, and male subjects belong to majority with 254 participants. Method also threw up two out of 100 respondents identified within the "Other" category.

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

Table 2: Demographic Profile of the Respondents

1. Age	Frequency
a. 18-30	48
b. 31-40	53
c. 41-50	128
d. 51-60	45
e. 60+	39
Total	313
2. Gender	
a. Male	254
b. Female	59
c. Other	0
Total	313
3. Position	
a. Financial Manager	56
b. Auditor	66
c. Accountant	92
d. Department Head	41
e. Other	58
Total	313
4. Years of Experience in the Banking Sector	
0-5 Years	58
6-10 Years	134
c. 11-15 Years	61
16-20 Years	45
20+ Years	15
Total	313

With respect to professional roles in the bank, accountants were the largest occupation with 92 respondents. Then were the high number of auditors (66), financial managers (56) and department heads (41).

Finally, 58 belong to different other professional roles of Banking sector. As for professional experience, 134 respondents reported having worked for between 6 and 10 years in the banking sector. Next came 61 respondents who have worked for between 11–15 years, and then there were 58 respondents who had less than six (6) to zero (0) years of experience. Further, 45 participants had experience between 16 & 20 years and only 15 respondents had more than twenty years of experience in the banking industry. Taken together, these demographic variables reveal that participants were mainly experienced banks with relevant occupational exposure to audit, finance and organisational risk management methods.

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

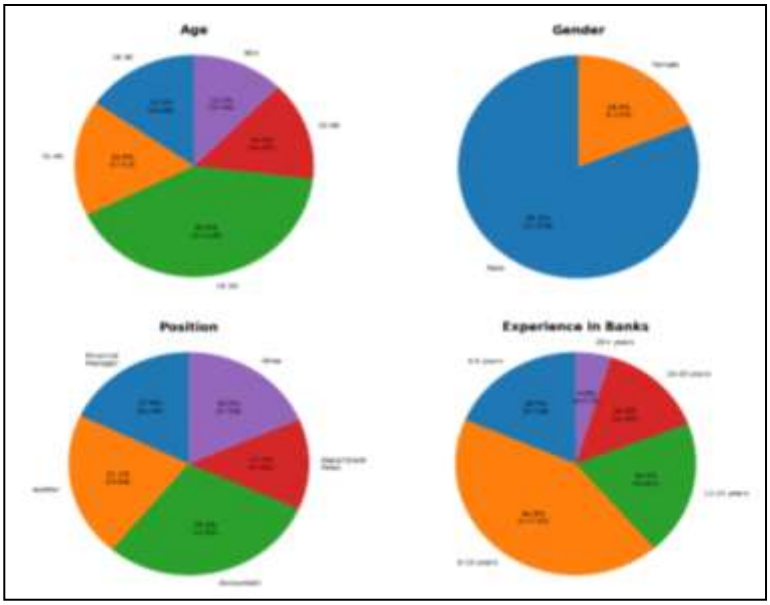

Figure 2: Demographic Factors of the Respondents

RESULTS AND DISCUSSION

As shown in Table 3, the results for reliability and validity assessment of five main constructs, namely AIRMF (AFR), AIRM, CFR, EDF as well as INA. As shown in Table 2 of statistical results, all constructs scored above the lowest acceptable threshold value (0.70), demonstrating adequate levels of internal consistency and reliability measurement for items reflecting the construct by having a Cronbach's alpha coefficient. The Cronbach's alpha coefficient of the AFR construct was 0.735, suggesting adequate internal consistency among its measurement items. Moreover, since the AVE value is 0.552, it indicates that an acceptable amount of variance relative to measurement error can be explained by the construct and confirms adequate convergent validity.

AIRM exhibited significantly better predictive power among the tested constructs, with a composite reliability value of 0.881 and an AVE value of 0.752 respectively. These results show good internal consistency and high convergent validity that the construct is closely related to the concept being measured as it consists of specific conceptual dimensions regarding AI-oriented risk governance practices. Likewise, CFR was marked by outstanding reliability characteristics. Internal consistency to measure this construct was high compared to other variables (Cronbach's alpha = 0.894, composite reliability score = 0.927). Summary of Findings Statistical indicators generally provide compelling evidence of high levels of reliability and acceptable levels of internal consistency for constituent measurement items relating to the framework.

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

Despite showing relatively weaker results than the best performing constructs, the EDF construct still achieved reliability and validity indicators within an acceptable range of methodological standards. In particular, the construct exhibited a Cronbach's alpha of 0.753 and an AVE value of 0.575 thus affirming sufficient internal consistency and satisfactory convergent validity for further statistical examination. The strongest combined output measurement was from the INA construct within the model. In terms of the construct, it was achieved a Cronbach's alpha of 0.908 and a composite reliability score of 0.944 that illustrates an extremely high reliability and considerable consistency among measurement indicators. These findings reaffirm the measurement structure of the INA construct is extremely stable and reliable. The statistical evidence collected is general and it emphasizes that all constructs of the study fulfilled established criteria in terms of reliability and validity assessment (as detailed in Table 3). Thus, the measurement model exhibits a suitable level of robustness, consistency, and construct adequacy and facilitates further inferential and structural analyses in this study.

Table 3: Reliability and Validity Investigation

	Cronbach's Alpha	Composite Reliability (rho_a)	Composite Reliability (rho_c)	Average Variance Extracted (AVE)
AFR	0.735	0.829	0.824	0.552
AIRM	0.772	0.881	0.858	0.752
CFR	0.894	0.897	0.927	0.760
EDF	0.753	0.762	0.843	0.575
INA	0.908	0.944	0.934	0.781

Note: AFR AIRMF Framework in Risk Mitigation, AIRM: AI Risk Management Strategies, CFR: COSO Framework, EDF: Evaluation of Deepfake, INA: Internal Auditing

Factor loading values in relation to selected measurement items for the study constructs are shown in Figure 3. The results of the assessment in SPSS show that the loading coefficients for all individual indicators are greater than 0.50 (minimum threshold value). These results show that for both constructs, each of the measurement items exhibits a sufficient level of association with their corresponding latent construct. The acceptable loadings values further validate that the observed items measured what they were intended to and represent the theoretical constructs. Thus, the results give empirical evidence for the construct validity of the measurement model on an individual-item level. From these statistical results, it can hence be concluded that the study has reached an acceptable level of validity with respect to the selected variables; thus, proving the measurement items being appropriate for structural analysis.

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

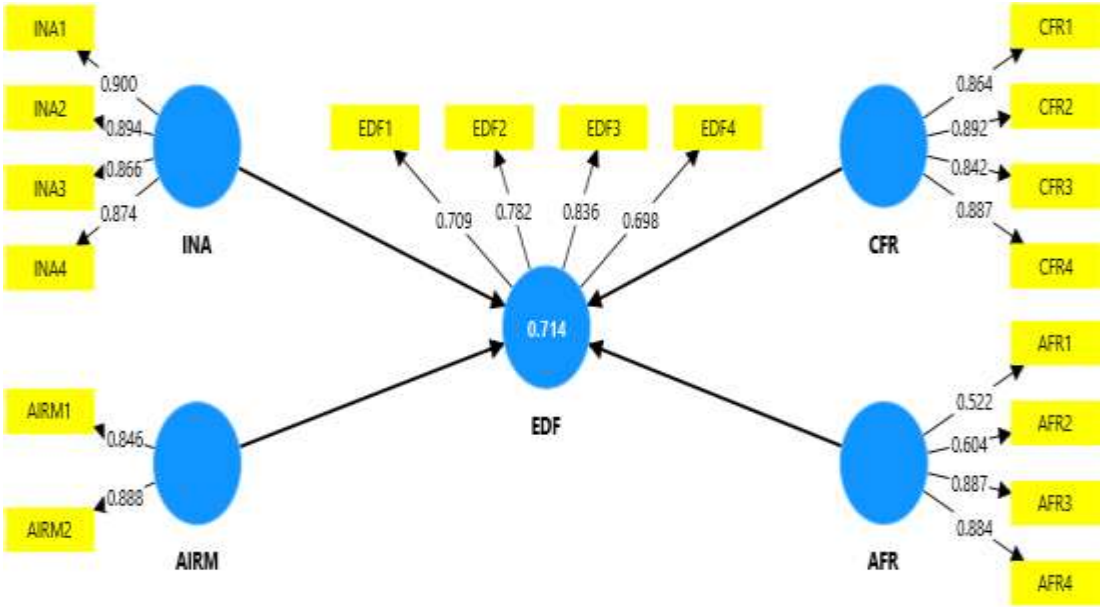

Figure 3: Loadings of the Items being Selected

Heterotrait–Monotrait Ratio (HTMT) values used to assess discriminant validity across the five main constructs in this study AFTR, AIRMM, CRF, EDF and INA have been noted in Table 4. The statistical results indicated that all HTMT coefficients are smaller than the established threshold value of 0.85, thus confirming that within the present study constructs are empirically distinct from each other. The highest value of HTMT among the relationships identified from the data with a coefficient of 0.809 between AFR and EDF. While this implication indicates a solid association associated with the general correlation of the alternatives, it does not exceed 0·85 and is thus indicative there are no major concerns about discriminant validity or too much overlap in their concepts.

Several construct relationships, however, were found to a be comparatively weaker. For example, the HTMT coefficient between AFR and CFR was 0.423, while the relationship between AIRM and CFR achieved a value of 0.442. Analogously, the relationship between EDF and INA provides a lower coefficient of 0.299. These somewhat smaller values indicate that the constructs are still conceptually distinct and act as separate dimensions in the measurement system. In summary, the HTMT results shown in Table 4 provide further solid empirical validity information for the discriminant validity theory of the measurement model. Results confirmed that constructs do measure theoretically distinct domains and provided enough discrimination between study variables for further structural and inferential analysis.

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

Table 4: HTMT Ratios

HTMT	AFR	AIRM	CFR	EDF	INA
AFR					
AIRM	0.541				
CFR	0.423	0.442			
EDF	0.809	0.339	0.573		
INA	0.274	0.284	0.197	0.299	

The Fornell–Larcker results are shown in Table 5, where the square roots of the AVE for each construct appear on the diagonal values, while off-diagonal values are correlations between constructs. The square root of AVE shows that how much a construct can (co)explain variance in its own indicators. Discriminant validity is the extent to which constructs are truly distinct from each other and is established when the square root of AVE is larger than correlations between that construct and the other latent constructs. In Table 5, the diagonal values (such as 0.743 for AFR and 0.867 for AIRM) are larger than the off-diagonal correlation values in the same group. For case in point, AIRM reported 0.867 but its correlations with other constructs (i.e., AFR = 0.408 and CFR = 0.347) were significantly lower. As a result, these findings validate discriminant validity in the research model.

Table 5: Fornell Larcker

Variables	AFR	AIRM	CFR	EDF	INA
AFR	0.743				
AIRM	0.408	0.867			
CFR	0.688	0.347	0.872		
EDF	0.664	0.746	0.475	0.758	
INA	0.262	0.229	0.188	0.260	0.883

As seen from Table 6, VIF values for selected measurement items. We use VIF to check for multicollinearity of our variables, when values greater than 5 or 10 is a characteristic of Multicollinearity. The results show that most of the measurement items had VIF values significantly below the threshold value of 5, which could be interpreted as low multicollinearity and therefore provides limited reason for concern. Some variables, particularly INA3 (3.674) and INA4 (3.894), were associated with higher VIF number than others, but still below the unacceptable threshold of 5. Thus, these results hint at the fact that multicollinearity is not a serious problem for current data. The structural model analysis indicates how much each of those explanatory constructs explains EDF. Discussion of the individual effects of the independent variables on EDF First, according to Table 6, it is positive through whom as AFR has a path coefficient of 0.453 and a p-value of 0.000, which means significant influence on the EDF variable. This result indicates that the AIRMF framework greatly aids in risk mitigation, improving deepfake evaluation application.

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

Table 6: VIF Scores of the Selected Items

	VIF
AFR1	1.491
AFR2	1.557
AFR3	2.638
AFR4	2.644
AIRM1	1.343
AIRM2	1.343
CFR1	2.590
CFR2	3.204
CFR3	2.360
CFR4	3.151
EDF1	1.244
EDF2	1.626
EDF3	1.833
EDF4	1.402
INA1	2.889
INA2	3.102
INA3	3.674
INA4	3.894

Essentially, by implementing AFR effectively, organisations can respond and mitigate potential threats posed by deepfake technology quicker and more efficiently. Consequently, AFR gives strong backing for the risk assessment of deepfakes. Moreover, Table 6 also shows that AIRM is found to be a crucial contributor and positively influences EDF by having coefficient value of (0.571) with p-value (0.000). This result suggests that AIRM is the most important construct influencing EDF among all constructs studied. Such a finding implies that organisations employing the right AI-based risk management strategies are more likely to identify, assess and respond to threats associated with deepfake technologies. Thus, AIRM is critical for risk management associated with AI-based systems and deepfake solutions.

On the contrary, estimates reported in Table 7 show that CFR has an adverse and statistically insignificant impact on EDF (coefficient value = -0.038; p-value = 0.390). This finding suggests that CFR has no practical effects on EDF in this study. While CFR will undoubtedly continue to have an important role in broader organisational control and governance practices around risk, these findings indicate that CFR's contribution toward more bespoke deepfake-specific risk assessment is somewhat limited. As a result, CFR does not lend significant support for EDF as compared to the other constructs. Moreover, INA also has a positive correlation with EDF (coefficient: 0.118; p-value: 0.000). These results provide evidence that INA has a positive impact on EDF, but relatively smaller than those of AFR and AIRM. On a practical level, internal auditors provide value to organisations by assessing risk exposures, as well as reviewing internal controls and processes that need to be in place to mitigate threats related to

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

unverified deepfakes. In this regard, INA still acts as an away player to improve the assessment of the risks exposed by deepfake.

Table 7: Structural Model Analysis

	Original Sample (O)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	P Values
AFR -> EDF	0.453	0.060	7.519	0.000
AIRM -> EDF	0.571	0.044	13.078	0.000
CFR -> EDF	-0.038	0.044	0.860	0.390
INA -> EDF	0.118	0.036	3.278	0.000

Figure 4 presents the Structural Equation Modelling (SEM) output associated with the proposed study framework.

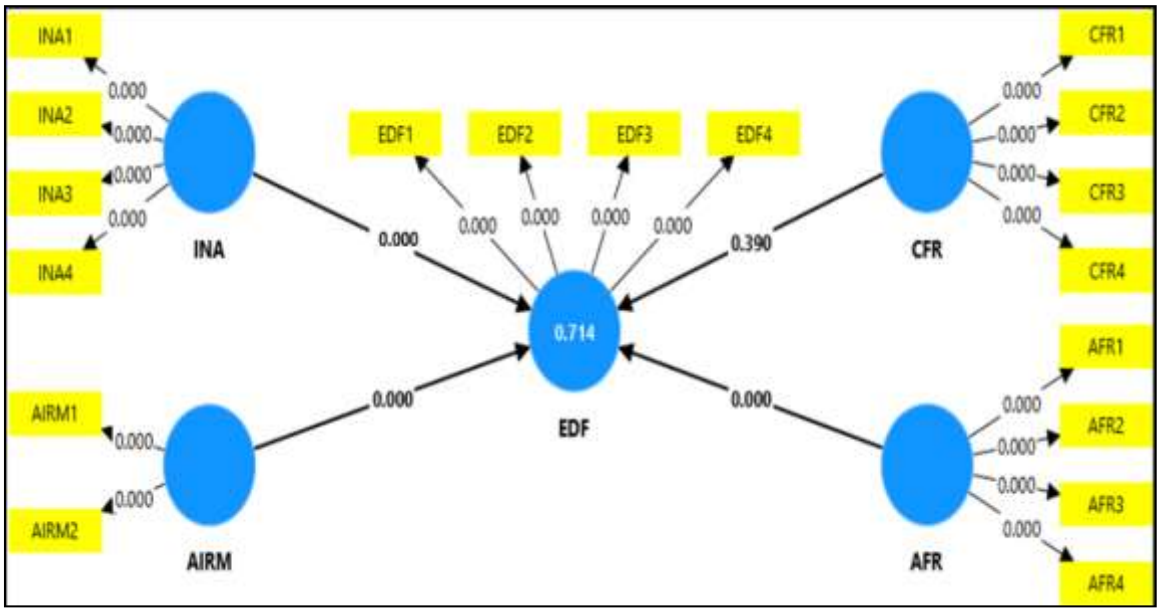

Figure 4: SEM Output

CONCLUSIONS AND RECOMMENDATIONS

The purpose of this study is to investigate the determinants of EDF in a developing economy, specifically focusing on the context of Iraq and providing new evidence within an under-researched field. This aim was investigated using four primary constructs: AFR, AIRM, CFR and INA. The data collected from 313 respondents were analysed and findings showed that AIRM and AFR had significant positive impact on EDF. On the other hand, INA also showed a beneficial contribution but to a lesser extent. On the other hand, CFR had shown no significant effects, implying that COSO might not be yet potent enough to meet with the AI-driven threats like Deepfakes in the

banking ecosystem. This reflection underscores the necessity for AIRM strategies that will empower banking organisations to perforate deepfake technologies more comprehensively. The data also highlights significant opportunities to bolster institutional resilience to new sources of AI-related risk although the Iraqi banking sector still faces numerous challenges relating to technology and governance.

The current study aims to detect the determinants of EDF in the Iraq banking sector. Purpose of the Study: In fulfilling this objective, the following four constructs were assessed: AFR, AIRM, CFR and INA. Based on the 313 responses, results showed evidence that AIRM and AFR had statistically significant positive impacts on EDF. Meanwhile, the accretive contribution from INA was less significant. A statistically significant effect was not found for CFR, however, so the COSO framework in its current form may be insufficient to address the complexities of artificial intelligence existence and deepfake challenges in banking ecology. These findings emphasize the required nature of implementing AIRM strategies for enabling banks to address and manage risks from deepfake technologies. Although the landscape of Iraq's banking sector is still plagued by significant Government and IT digital merit challenges, these results shed light on key opportunities to strengthen institutional resilience in preparation for AI-led threats soon.

Setting up a solid AIRM framework is thus a strategic imperative to Iraqi banks. With improvements to such frameworks, banks can better identify, assess and manage risks stemming from AI systems including those that might result from deepfake technology. Accordingly, banks are urged to provide extensive AI governance policies and systems along with in-house monitoring mechanisms as well as comprehensive artificial intelligence guidelines (AI Guidelines) that will foster and strengthen operational trust, security and technology reliability within their systems. Lastly, while INA had a somewhat diminished effect on EDF relative to the other dimensions, this variable still played an important positive role in terms of governance overall. Thus, these operations of INA must be improved further in the Iraqi banks through more focus on reviewing AI systems and validating risk control/assessment and cybersecurity measures. Audit reviews, assessment of compliance would further help unearth vulnerabilities that are baked into AI embedded operational environments.

The results also indicated that CFR do not have a relatively important role in influencing EDF which necessitates that governance in the banking sector has to be strengthened toward specialised and artificial intelligence. As a result, it may require near term alterations to traditional COSO based frameworks that banking institutions use, augmenting them even further with the unique challenges unleashed due to recent innovations in AI technologies. For example, it might include newer risk assessment methodologies and specific governance mechanisms tied to AI that can respond adequately to the new class of risks related to deep fakes. The paper also highlights that, to mitigate the AI-related risks, particularly those posed by deep fake capabilities,

collaboration among Iraqi banks and government regulators or technologists is critical. Capacity building and material development to build partnerships for knowledge pooling keeps banks updated with newer threats, evolving methods of technology and standards of AI governance while the digital market continues to be dynamic.

REFERENCES

Adavelli, S. R., Mittapelly, A. K., & reddy Karri, N. (2025). AI and Cybersecurity: Advancements in Threat Detection and Prevention. Academic Guru Publishing House.

Arif, A., Khan, M. I., & Khan, A. R. A. (2024). An overview of cyber threats generated by AI. International Journal of Multidisciplinary Sciences and Arts, 3(4), 67-76 DOI: https://doi.org/10.47709/ijmdsa.v3i4.4753.

Bharati, R. K. (2024). AI-enhanced social engineering: evolving tactics in cyber fraud and manipulation. The Academic–International Journal of Multidisciplinary Research (A Peer Reviewed Refereed Online Journal), 2(7), 16-26.

Chang, S.-I., Huang, S.-M., Roan, J., Chang, I. C., & Liu, P.-J. (2014). Developing a risk management assessment framework for public administration in Taiwan. Risk Management, 16(3), 164-194. https://doi.org/10.1057/rm.2014.9

Chen, H., & Magramo, K. (2024). Finance worker pays out $25 million after video call with deepfake 'chief financial officer'. CNN, February, 4.

Chen, M. (2024). Finance worker pays out $25 million after video call with deepfake 'chief financial officer'. CNN, February, 4. . https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk

Coetzee, P. (2016). Contribution of internal auditing to risk management: Perceptions of public sector senior management. International Journal of Public Sector Management, 29(4), 348-364. https://doi.org/10.1108/IJPSM-12-2015-0215

Das, R., Ahmed, W., Sharma, K., Hardey, M., Dwivedi, Y. K., Zhang, Z., Apostolidis, C., & Filieri, R. (2024). Towards the development of an explainable e-commerce fake review index: An attribute analytics approach. European Journal of Operational Research, 317(2), 382-400. https://doi.org/10.1016/j.ejor.2024.03.008

De Bock, K. W., Coussement, K., Caigny, A. D., Słowiński, R., Baesens, B., Boute, R. N., Choi, T.-M., Delen, D., Kraus, M., Lessmann, S., Maldonado, S., Martens, D., Óskarsdóttir, M., Vairetti, C., Verbeke, W., & Weber, R. (2024). Explainable AI for Operational Research: A defining framework, methods, applications, and a research agenda. European Journal of Operational Research, 317(2), 249-272. https://doi.org/10.1016/j.ejor.2023.09.026

Farid, H. (2022). Creating, using, misusing, and detecting deep fakes. Journal of Online Trust and Safety, 1(4), 1-33. https://doi.org/10.54501/jots.v1i4.56

Filieri, R., Lin, Z., Li, Y., Lu, X., & Yang, X. (2022). Customer Emotions in Service

Robot Encounters: A Hybrid Machine-Human Intelligence Approach. Journal of Service Research, 25(4). https://doi.org/10.1177/10946705221103937

Gambín, Á. F., Yazidi, A., Vasilakos, A., Haugerud, H., & Djenouri, Y. (2024). Deepfakes: current and future trends. Artificial Intelligence Review, 57(3), 64. https://doi.org/10.1007/s10462-023-10679-x

Güera, D., & Delp, E. J. (2018). Deepfake video detection using recurrent neural networks. 2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS), https://doi.org/10.1109/AVSS.2018.8639163

Haris, L. (2020). CBS News asks Facebook to remove "deepfake" video of Mark Zuckerberg with unauthorized CBSN trademark. https://www.cbsnews.com/news/cbs-news-asks-facebook-to-remove-deep-fake-video-of-mark-zuckerberg-with-unauthorized-cbsn-trademark/

Islam, M. B. E., Haseeb, M., Batool, H., Ahtasham, N., & Muhammad, Z. (2024). AI threats to politics, elections, and democracy: A blockchain-based deepfake authenticity verification framework. Blockchains, 2(4), 458-481. https://doi.org/10.3390/blockchains2040020

Jayashre, K., & Amsaprabhaa, M. (2024). Safeguarding media integrity: A hybrid optimized deep feature fusion based deepfake detection in videos. Computers & Security, 142, 103860. https://doi.org/10.1016/j.cose.2024.103860

Kavanagh, C. (2022). New Tech, New Threats, and New Governance Challenges: An Opportunity to Craft Smarter Responses? Carnegie Endowment for International Peace. https://assets.production.carnegie.fusionary.io/static/files/files__WP_Camino_Kavanagh___New_Tech_New_Threats1.pdf

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? Business horizons, 63(2), 135-146. https://doi.org/10.1016/j.bushor.2019.11.006

Mandal, S. (2025). Deep Fake Technology and Identity Theft: An Emerging Challenge for Cyber Laws in India. Available at SSRN 5161545. https://dx.doi.org/10.2139/ssrn.5161545

Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A., & Malik, H. (2023). Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward: Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward. Applied intelligence, 53(4), 3974-4026. https://doi.org/10.1007/s10489-022-03766-z

Nawaz, M., Javed, A., & Irtaza, A. (2024). A deep learning model for FaceSwap and face-reenactment deepfakes detection. Applied Soft Computing, 162, 111854. https://doi.org/10.1016/j.asoc.2024.111854

Perot, E., & Mostert, F. (2020). Fake it till you make it: an examination of the US and English approaches to persona protection as applied to deepfakes on social media. Journal of Intellectual Property Law & Practice, 15(1), 32-39. https://doi.org/10.1093/jiplp/jpz164

Vol: 18 **No:** 02 **Year:** 2026 **ISSN:** 1309-8055 (Online) (**pp.** 162-180) **Doi:** 10.34109/ijefs. 2026180207

Portuguez-Castro, M. (2025). of Technological Disruption. Technology and Society-Boon or Bane?: Business Systems Laboratory International Symposium, Varese, Italy, 2025, https://iris.unipa.it/handle/10447/700256

Rajput, T., & Arora, B. (2023). A Systematic review of deepfake detection using learning Techniques and Vision Transformer. International Conference on Cognitive Computing and Cyber Physical Systems, https://doi.org/10.1007/978-981-97-2550-2_17

Rehaan, M., Kaur, N., & Kingra, S. (2024). Face manipulated deepfake generation and recognition approaches: A survey. Smart Science, 12(1), 53-73. https://doi.org/10.1080/23080477.2023.2268380

Sai, S., & Wang, Z. (2026). Criminal Regulatory Approaches to Deepfake-Related Offenses: Focusing on the Crime of Fraud. International Journal of Asian Social Science Research, 3(1), 20-31. https://doi.org/10.70267/ijassr.260301.2031

Tariq, S., Jeon, S., & Woo, S. S. (2022). Am I a real or fake celebrity? Evaluating face recognition and verification APIs under deepfake impersonation attack. Proceedings of the ACM Web Conference 2022, https://doi.org/10.1145/3485447.35122

Usmani, S., Kumar, S., & Sadhya, D. (2026). Deepfakes: A comprehensive survey on techniques, challenges and future directions. Computers and Electrical Engineering, 136, 111210. https://doi.org/10.1016/j.compeleceng.2026.111210

Vecchietti, G., Liyanaarachchi, G., & Viglia, G. (2025). Managing deepfakes with artificial intelligence: Introducing the business privacy calculus. Journal of Business Research, 186, 115010. https://doi.org/https://doi.org/10.1016/j.jbusres.2024.115010

Walczyna, T., & Piotrowski, Z. (2023). Quick overview of face swap deep fakes. Applied Sciences, 13(11), 6711. https://doi.org/10.3390/app13116711