
AI IN AUDITING AND FRAUD DETECTION

***¹Komal Kumari, ¹Kengeri Chaithra P, ²Dr. Anusha. K. P.**

¹Master of Business Administration, BGS B SCHOOL, Kengeri.²Assistant Professor, BGS B School.

Article Received: 27 August 2026, Article Revised: 17 September 2026, Published on: 07 October 2026

***Corresponding Author: Komal Kumari**

Master of Business Administration, BGS B SCHOOL, Kengeri.

Doi: <https://doi-doi.org/101555/ijarp.5007>**ABSTRACT**

The increasing use of Artificial Intelligence (AI) in auditing is changing the way financial information is examined and potential fraudulent activities are identified. AI-based techniques can process extensive financial datasets and assist auditors in recognizing unusual transactions, recurring patterns, and potential indicators of fraud. Despite these advantages, the limited interpretability of many AI models creates challenges for auditors who are responsible for evaluating and supporting their professional decisions.

This study focuses on Explainable Artificial Intelligence (XAI) as a means of improving the transparency of AI-assisted fraud detection. Although previous research has largely concentrated on the predictive capability and adoption of AI-based auditing tools, comparatively less attention has been given to the behavioural and professional factors associated with explainability. In particular, the relationships among explainability, perceived transparency, auditor trust, professional scepticism, and fraud-detection effectiveness require further consideration.

The study uses a qualitative secondary-data approach based on academic publications, professional auditing standards, industry reports, and relevant regulatory documents. The literature is reviewed and synthesized to identify major themes, relationships, limitations, and opportunities for future research.

Based on the reviewed literature, the study proposes a conceptual framework in which XAI improves the interpretability of AI-generated fraud alerts, thereby increasing perceived transparency and supporting appropriate auditor trust. Professional scepticism is positioned as an important safeguard that encourages auditors to critically evaluate AI recommendations rather than accept them automatically. The framework therefore emphasizes collaboration

between AI systems and human auditors, with AI supporting analytical tasks while final professional responsibility remains with the auditor.

KEYWORDS: Artificial Intelligence (AI), Explainable Artificial Intelligence (XAI), Fraud Detection, Auditing, Auditor Judgment, Professional Scepticism, Perceived Transparency, Human–AI Collaboration, Secondary Data, Literature Review.

INTRODUCTION

Artificial Intelligence has become increasingly relevant to the auditing profession as organizations generate larger and more complex volumes of financial and transactional information. Conventional audit procedures may face difficulties when potential irregularities are hidden within extensive datasets. AI-based approaches, including machine learning, deep learning, and anomaly-detection techniques, can assist auditors by examining large populations of transactions and highlighting unusual patterns that may require further investigation.

The use of AI, however, also introduces a significant challenge: auditors may not always be able to understand how an AI system arrived at a particular prediction or fraud alert. Such systems are frequently described as “black boxes” because their internal decision processes may be difficult for users to interpret. This creates concerns about transparency and accountability, particularly because auditors remain responsible for evaluating evidence and forming professional conclusions.

Explainable Artificial Intelligence provides an approach for addressing this limitation. XAI techniques can provide information about the factors contributing to an AI prediction, allowing auditors to examine the reasoning associated with a fraud alert. Greater interpretability may therefore support informed use of AI, improve confidence in appropriate situations, and encourage auditors to maintain professional scepticism.

Although considerable research has examined the technical performance and adoption of AI-based auditing systems, the behavioural and professional consequences of explainability remain comparatively less developed. In particular, greater attention is required to understand how XAI may influence perceived transparency, auditor trust, professional scepticism, and the effectiveness of fraud detection.

Accordingly, this study examines the role of Explainable Artificial Intelligence in AI-assisted fraud detection within auditing. It develops a conceptual framework connecting XAI with perceived transparency, auditor trust, professional scepticism, and fraud-detection

effectiveness. The study aims to provide a foundation for future empirical investigation and to offer practical insights for auditors, audit firms, regulators, and policymakers.

LITERATURE REVIEW

1. Zhang, Cho, and Vasarhelyi (2022) – Explainable Artificial Intelligence (XAI) in Auditing : Zhang, Cho, and Vasarhelyi (2022) emphasized the importance of interpretability in AI-supported auditing. Their work suggests that explanations accompanying AI-generated outputs can help auditors understand automated recommendations and assess them within the context of audit evidence. Techniques such as SHAP, LIME, and rule-based explanations can contribute to documentation and professional judgement. The study supports the view that XAI should strengthen auditor expertise rather than substitute for it.

2. Tümmeler and Quick (2025) – How to Detect Fraud in an Audit: A Systematic Review of Experimental Literature: Tümmeler and Quick (2025) reviewed experimental research on fraud detection and showed that audit effectiveness depends on both technological support and human factors. Professional scepticism, fraud-risk assessment, auditor experience, and decision-support systems remain important to the detection process. Their findings indicate that AI can assist in identifying suspicious activity, while auditors must continue to evaluate the results critically.

3. Zafar and Wu (2026) – Methodological Challenges in Explainable AI for Fraud Detection: Zafar and Wu (2026) reviewed research on explainable AI and identified methodological concerns including class imbalance, inconsistent validation, and difficulties in evaluating explanation quality. Their review also highlighted the widespread use of SHAP and LIME. The findings indicate that a highly accurate prediction does not necessarily provide a sufficiently useful or reliable explanation, creating a need for stronger evaluation approaches.

4. Faccia (2026) – Explainable AI (XAI) in Auditing: Bridging the Gap Between Predictive Fraud Models and Regulatory Standards: Faccia (2026) considered XAI from regulatory and practical perspectives and argued that the development of AI-assisted auditing should balance predictive capability with transparency and accountability. The study connects explainability with governance, ethical considerations, auditor confidence, and regulatory expectations, suggesting that interpretable systems will become increasingly relevant to responsible audit practice.

5. Arias-González et al. (2026) – Explainable Artificial Intelligence in Accounting and Financial Auditing: A Systematic Review: Arias-González et al. (2026) reviewed applications of XAI in accounting and financial auditing, including fraud detection, risk assessment, compliance, and decision support. Their findings identify SHAP and LIME as frequently used techniques while also noting practical challenges involving data quality, computational demands, explanation stability, and regulatory adaptation. The review points toward the need for stronger explanation methods and closer integration of human expertise with AI.

RESEARCH GAP

The existing literature demonstrates considerable progress in applying AI to fraud detection and auditing. However, much of the research continues to emphasize technological performance, algorithm development, anomaly identification, and predictive accuracy. Comparatively less attention has been given to how auditors interpret and respond to AI-generated outputs.

A further limitation is the fragmented treatment of technological and behavioural factors. Explainability, perceived transparency, auditor trust, and professional scepticism are often examined independently rather than as interconnected elements of AI-assisted fraud detection. Consequently, there is scope for a framework that considers how these factors interact and contribute to effective human–AI collaboration in auditing.

OBJECTIVES

1. To examine the contribution of Artificial Intelligence to fraud detection and contemporary auditing practices.
2. To identify the transparency-related limitations associated with black-box AI fraud-detection systems.
3. To examine how Explainable Artificial Intelligence can improve auditors' interpretation of AI-generated fraud alerts.
4. To analyze the relationship between XAI and auditor trust in AI-supported fraud detection.

RESEARCH METHODOLOGY

1. Research Approach

The study adopts a qualitative research approach to examine how Explainable Artificial Intelligence supports auditors in interpreting and evaluating AI-generated fraud-detection outputs.

2. Research Design

A descriptive and exploratory design is used. The descriptive component outlines the role of AI and XAI in auditing, while the exploratory component considers the relationships among explainability, auditor trust, professional scepticism, accountability, and fraud-detection effectiveness.

3. Data Source

The study relies exclusively on secondary information obtained from peer-reviewed journals, research articles and scholarly publications, books and academic databases, professional auditing standards and guidelines, reports from accounting and auditing bodies, industry reports, regulatory documents, and published material concerning AI, XAI, fraud detection, and auditing.

4. Data Collection Method

A systematic literature review and document-analysis approach is adopted. Relevant sources concerning AI-based fraud detection and XAI in auditing are identified, reviewed, compared, and synthesized to establish existing knowledge, major themes, and research gaps.

5. Data Analysis

The secondary material is examined through thematic analysis. The review focuses on recurring themes related to explainability, transparency, auditor trust, professional scepticism, human–AI interaction, accountability, fraud-detection effectiveness, and ethical AI governance. These themes are compared and synthesized to support development of the proposed conceptual framework.

6. Scope of the Study

The study focuses on the application of Explainable Artificial Intelligence to AI-assisted fraud detection in auditing. Particular attention is given to its potential contribution to auditor understanding, trust, professional judgement, professional scepticism, accountability, and effective human–AI collaboration.

7. LIMITATIONS

The study is conceptual and is based exclusively on secondary sources; therefore, the proposed relationships have not been empirically tested using primary data. The conclusions may also be influenced by the availability, scope, and quality of existing literature. In

addition, the framework may not capture differences in AI adoption across countries, audit firms, industries, or regulatory environments.

Future research can address these limitations through surveys or interviews with audit professionals, cross-country comparisons, longitudinal studies, and empirical testing of additional factors such as AI literacy, organizational support, regulatory compliance, and auditor experience. Future studies may also compare explainable and conventional AI systems and examine industry-specific applications of XAI.

EXPECTED OUTCOME

The study is expected to indicate that greater explainability can improve the usefulness of AI-assisted fraud detection by enabling auditors to understand the basis of AI-generated alerts. Improved transparency may contribute to appropriate levels of auditor trust, while professional scepticism can help prevent unquestioning acceptance of automated recommendations. Overall, the proposed framework is expected to support the view that effective fraud detection requires a balanced combination of AI-based analytical capability and human professional judgement.

DISCUSSION

The reviewed literature indicates that the usefulness of AI in fraud detection depends not only on its ability to identify unusual patterns but also on the extent to which auditors can interpret and evaluate its outputs. When an AI system produces a fraud alert without providing understandable reasoning, the auditor may find it difficult to assess whether the recommendation is supported by sufficient evidence.

XAI can address this issue by making important aspects of an AI prediction more accessible to the auditor. When the factors contributing to a particular alert are visible, auditors can investigate the underlying transactions, compare the AI recommendation with other audit evidence, and determine whether additional procedures are necessary. Thus, explainability can strengthen the role of AI as a decision-support mechanism rather than positioning it as an independent decision-maker.

The literature also suggests a connection between interpretability and auditor trust. An auditor who understands why an AI system has generated a particular recommendation may be better positioned to judge when that recommendation is useful and when it requires further verification. Appropriate trust therefore differs from unquestioning reliance: the purpose of explainability is to support informed evaluation rather than automatic acceptance.

Professional scepticism provides an important control within this relationship. Auditors are expected to critically evaluate evidence and remain alert to the possibility of error, bias, or misleading information. In an AI-supported environment, this responsibility remains relevant because an apparently accurate prediction may still be based on incomplete, biased, or inappropriate data. Explainability can therefore encourage auditors to investigate the basis of an AI-generated result and seek supporting evidence before reaching a final conclusion.

The proposed framework consequently views human and artificial intelligence as complementary. AI can assist with computationally intensive activities such as identifying unusual transactions and patterns, while auditors contribute contextual understanding, professional judgement, ethical reasoning, and responsibility for the final audit decision.

The review also points to the importance of governance. Transparent AI systems can make automated decisions easier to document, review, and challenge. Nevertheless, explainability alone cannot guarantee responsible AI use. Factors such as data quality, auditor competence, organizational readiness, system reliability, privacy, and regulatory requirements also influence successful implementation.

Overall, the literature supports the potential of XAI to strengthen AI-assisted fraud detection, but empirical evidence is still required to establish the strength of the proposed relationships. Future research should therefore test whether explainability actually improves auditor judgement, trust calibration, professional scepticism, and fraud-detection outcomes in practical audit settings.

PRACTICAL IMPLICATIONS

The study has implications for audit firms, auditors, AI developers, and regulatory authorities. For audit firms, the adoption of XAI-enabled fraud-detection tools can improve the interpretability of automated alerts and assist auditors in deciding which transactions require closer examination.

Auditor training is also important. Professionals need sufficient AI literacy to understand the capabilities and limitations of AI systems and to evaluate explanations critically. Training should therefore focus not only on operating AI tools but also on maintaining professional scepticism when interpreting their outputs.

For AI developers, the findings highlight the importance of designing systems that communicate their reasoning in a form that is understandable and useful to audit professionals. Explainability should be incorporated into system design rather than treated as an additional feature after the model has been developed.

Regulators and professional accounting bodies can also contribute by establishing guidance concerning transparency, documentation, accountability, data governance, and the responsible use of AI in auditing. Such guidance could help organizations adopt AI while preserving the auditor's responsibility for professional decisions.

Finally, the proposed framework can serve as a starting point for empirical studies examining the relationships among explainability, transparency, trust, professional scepticism, and fraud-detection effectiveness.

CONCLUSION

This study highlights the potential of Explainable Artificial Intelligence to strengthen fraud detection within the auditing profession. The review suggests that the value of AI should not be assessed solely through predictive performance; auditors must also be able to understand, question, and evaluate the outputs generated by AI systems.

The proposed framework therefore places explainability, perceived transparency, auditor trust, and professional scepticism within a connected model of AI-assisted fraud detection. XAI may help auditors interpret automated recommendations, while professional scepticism ensures that such recommendations remain subject to critical human evaluation.

The study contributes a conceptual perspective on the role of human–AI collaboration in auditing and provides a basis for future empirical research. Further investigation using primary data from audit professionals is necessary to determine whether the relationships proposed in the framework are supported in actual audit environments.

REFERENCES

1. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
2. Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, 19(2), 213–236.
3. International Auditing and Assurance Standards Board (IAASB). (2024). International Standards on Auditing.
4. Rai, A. (2020). Explainable AI: From Black Box to Glass Box. *Journal of the Academy of Marketing Science*.
5. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425–478.
6. Zhang, Cho and Vasarhelyi (2022), Tümmeler and Quick (2025), Zafar and Wu (2026), Faccia (2026), and Arias-González et al. (2026).