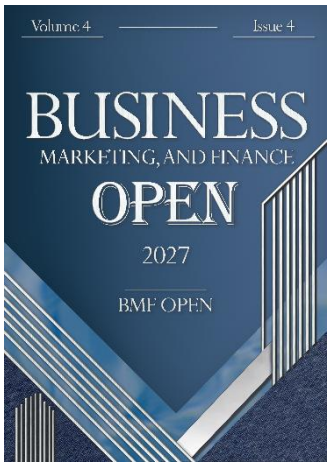
An Explainable Auditing Framework for Assessing Financial Statement Risk Based on Explainable Artificial Intelligence: A Mixed-Methods Approach

Rahim Mollaei[1], Ali Jafari[2], Asgar Pakmaram[3,*] and Nader Rezaei[4]



[1] PhD Student in Accounting, Department of Acconting, Bon.C., Islamic Azad University, Bonab, Iran;
[2] Department of Acconting, Bon.C., Islamic Azad University, Bonab, Iran;
[3] Department of Acconting, Bon.C., Islamic Azad University, Bonab, Iran;
[4] Department of Acconting, Bon.C., Islamic Azad University, Bonab, Iran;

* Correspondence e-mail: pakmaram@iau.ac.ir

Citation: Mollaei, R., Jafari, A., Pakmaram, A., & Rezaei, N. (2027). An Explainable Auditing Framework for Assessing Financial Statement Risk Based on Explainable Artificial Intelligence: A Mixed-Methods Approach. *Business, Marketing, and Finance Open*, 4(4), 1-29.

Received: 03 March 2026
Revised: 14 July 2026
Accepted: 22 July 2026
Initial Publication: 24 July 2026
Final Publication: 01 July 2027

Abstract: This study aimed to design and validate an explainable auditing framework for assessing financial statement risk based on explainable artificial intelligence. Grounded in the pragmatist paradigm, which enables the integration of inductive and deductive approaches, this study employed a mixed-methods methodology with a sequential exploratory design and was developmental–applied in terms of its objective. In the qualitative phase, 39 selected scientific sources were analyzed from an initial pool of 188 sources using Sandelowski and Barroso's seven-stage model. The coding process yielded 102 codes, which were subsequently organized into 42 initial codes, 18 organizing codes, and four overarching codes: "explainable auditing based on explainable artificial intelligence," "antecedents," "outcomes," and "mediating factors." The reliability of the qualitative analysis was confirmed using Cohen's kappa coefficient (0.742). In the quantitative phase, data were collected through a questionnaire administered to 98 auditors, accountants, and information technology specialists. The instrument's validity was confirmed using the content validity ratio (CVR) and content validity index (CVI), while its reliability was established using Cronbach's alpha. Partial least squares structural equation modeling (PLS-SEM) was employed to test the conceptual model. The findings indicate that organizational and infrastructural requirements, as well as professional trust-building, play significant roles in producing operational and economic outcomes and improving audit quality. Moreover, auditors' trust in explainable artificial intelligence models functions as a mediating factor affecting certain outcomes.

Keywords: explainable auditing, financial statement risk, explainable artificial intelligence, mixed-methods approach

1. Introduction

The rapid digitalization of business processes has substantially transformed the nature, volume, velocity, and complexity of financial information available to organizations, auditors, regulators, and other stakeholders. Contemporary financial reporting systems generate large quantities of structured and unstructured data that exceed the processing capacity of conventional audit procedures when they are examined exclusively through manual sampling, traditional analytical review, and rule-based professional judgment. At the same time, financial statement users increasingly expect auditors to identify material misstatements, fraud risks, anomalous transactions, and emerging patterns with greater speed, precision, consistency, and transparency. These developments have created strong incentives

for audit firms and financial institutions to adopt artificial intelligence, machine learning, robotic process automation, and advanced audit data analytics. Such technologies can assist auditors in examining entire populations of transactions, recognizing complex associations, prioritizing high-risk accounts, and producing more timely risk assessments [1-3]. Nevertheless, the value of algorithmic systems in auditing depends not only on predictive accuracy but also on whether auditors can understand, evaluate, challenge, document, and defend the reasoning underlying their outputs.

Financial statement risk assessment is one of the most consequential stages of the audit process because it shapes audit planning, the allocation of resources, the selection of substantive procedures, and the determination of the nature, timing, and extent of audit evidence. The assessment of financial risk requires the integration of accounting information, internal control characteristics, management behavior, industry conditions, transaction patterns, and contextual indicators. Traditional methods may be insufficient when risk signals are nonlinear, multidimensional, or distributed across large datasets. Machine-learning systems can improve the detection of unusual relationships and latent risk indicators in financial statements, while systematic evidence indicates that these models have become increasingly relevant for identifying misstatement and fraud risk [4-6]. However, a model that produces a highly accurate risk score without a comprehensible explanation may be difficult to incorporate into an audit engagement because the auditor must establish how the identified risk relates to assertions, accounts, transactions, controls, and the risk of material misstatement.

Artificial intelligence in auditing is therefore associated with a fundamental tension between predictive performance and professional interpretability. Complex models, including deep-learning architectures and ensemble techniques, can identify patterns that simpler models may overlook, but their internal decision processes may be opaque. In audit settings, opacity can weaken the auditor's ability to exercise professional skepticism, assess contradictory evidence, determine whether an output is reasonable, and explain the basis of a conclusion to reviewers or regulators. A technically sophisticated result may consequently have limited professional value when the pathway from data input to risk classification cannot be traced. Research on auditor judgment has shown that artificial intelligence can affect the intelligence, design, and choice stages of decision-making, but the quality of these effects depends on the way algorithmic outputs are integrated with professional reasoning rather than accepted mechanically [7-9]. This requirement has made explainable artificial intelligence increasingly important in the development of audit-support systems.

Explainable artificial intelligence refers to methods, models, and design principles intended to make the operation and outputs of artificial intelligence systems understandable to human users. Explainability may be intrinsic, as in decision trees and rule-based systems, or post hoc, where an explanation is generated for a complex model after the prediction has been produced. It may also be local, focusing on the explanation of a specific prediction, or global, clarifying the overall behavior of the model. The XAI literature emphasizes that explanation should not be treated as a single technical property, because the usefulness of an explanation depends on the user, decision context, risk level, institutional purpose, and type of model employed [10-12]. In auditing, an explanation must therefore do more than identify influential variables; it should connect the model's output to audit objectives, professional standards, materiality, relevant assertions, audit evidence, and the logic of risk assessment.

Among the most widely discussed post hoc explanation techniques are SHAP and LIME. SHAP estimates the contribution of each variable to a prediction using principles derived from cooperative game theory, whereas LIME approximates the behavior of a complex model around an individual observation using a simpler interpretable model. Both techniques can help auditors identify which financial ratios, transactions, account balances, or

behavioral indicators contributed to a high-risk classification [13]. Nevertheless, explanation methods may vary in stability, fidelity, consistency, and sensitivity to data changes. An explanation that changes substantially after a minor modification to the input may create uncertainty rather than clarity. Consequently, the adoption of XAI in auditing requires criteria for evaluating explanation quality, including accuracy, comprehensibility, relevance, consistency, actionability, and alignment with domain knowledge [14, 15]. These criteria are particularly important because auditors are accountable for the final professional judgment even when an artificial intelligence system contributes to the analysis.

The application of XAI in auditing can support several interconnected objectives. First, it can improve the traceability of the audit process by recording how data were processed, which variables were influential, and why specific accounts or transactions were classified as high risk. Second, it can strengthen the defensibility of audit decisions by providing an evidential link between algorithmic output and professional judgment. Third, it can improve the communication of risk assessments among engagement-team members, audit committees, regulators, and clients. Fourth, it can reduce the danger that auditors rely mechanically on automated outputs without critically examining their validity. Transparent AI applications may thus contribute to both audit efficiency and accountability, provided that explanations are designed according to the informational needs of professional users [16-18]. Explainability is therefore not merely a technical supplement to an audit model; it is a mechanism through which algorithmic analysis can become professionally usable and institutionally legitimate.

The concept of explainable auditing extends beyond the technical explanation of model outputs. It involves an integrated arrangement of data governance, model design, professional standards, human competencies, organizational processes, ethical safeguards, and accountability mechanisms. Explainability auditing has been proposed as a multidisciplinary process through which intelligent systems are examined not only for predictive capability but also for the appropriateness, reliability, accessibility, and social consequences of their explanations [19]. In regulated sectors, an effective framework should clarify what must be explained, to whom explanations should be provided, at which stage of the decision process they are needed, and how their adequacy should be evaluated [20, 21]. This broader perspective is especially relevant in external and internal auditing, where the consequences of erroneous risk classification may include inappropriate audit procedures, undetected material misstatement, excessive audit effort, regulatory exposure, and reputational damage.

The professional acceptance of explainable auditing is also shaped by auditors' trust in artificial intelligence models. Trust is not equivalent to unquestioning reliance. Calibrated trust requires auditors to believe that a system is sufficiently reliable and useful while maintaining the capacity to question its assumptions, limitations, and outputs. Auditor acceptance is likely to increase when the model is transparent, its outputs are stable, its explanations are relevant to audit objectives, and responsibility for the final decision remains clearly assigned to the auditor. Empirical studies have identified perceived usefulness, ease of understanding, technological readiness, professional competence, and institutional support as important factors influencing auditors' acceptance of XAI [22, 23]. Human-in-the-loop approaches are particularly important because they preserve the role of auditor expertise and enable the professional to confirm, revise, or reject an algorithmic recommendation [24]. Such arrangements help prevent both automation bias and excessive distrust of technology.

Organizational and human readiness represent another major condition for the successful implementation of explainable auditing. Audit firms require personnel who possess combined knowledge of accounting, auditing, data analytics, information systems, and artificial intelligence. Auditors must understand the practical implications of model outputs, while data specialists must recognize the professional and evidential requirements of auditing.

Cross-functional collaboration is therefore necessary to translate technical explanations into professionally meaningful interpretations. The sustainability of the audit profession in the digital era increasingly depends on digital competencies, the ability to detect fraud risk, and the continuous development of technological capabilities [25]. Organizational learning, management support, and applied training can further facilitate the adoption of advanced analytical tools [20]. The absence of these capabilities may lead to superficial use of XAI, misinterpretation of explanations, or the creation of an illusion of understanding.

The risks associated with explainable artificial intelligence must also be recognized. Explanations may appear persuasive without faithfully representing the true operation of the underlying model. Users may misinterpret feature importance as causality, overestimate the certainty of a prediction, or treat a simplified explanation as a complete account of the model. Research on causality in XAI emphasizes that predictive association should not automatically be interpreted as a causal relationship [26]. This distinction is critical in auditing because a variable associated with financial statement risk may not be the actual cause of the underlying misstatement. In addition, explanation systems may reproduce biased data patterns, normalize problematic assumptions, or transfer perceived responsibility from professionals to algorithms. These risks can be intensified when complex systems are adopted without adequate governance or when auditors lack sufficient knowledge to critically evaluate the evidence provided [27, 28].

Ethical and legal considerations are similarly central to the design of explainable auditing. The use of financial and behavioral data raises concerns related to privacy, confidentiality, fairness, data ownership, and unauthorized secondary use. Algorithmic decisions may affect the treatment of clients, the allocation of audit attention, and the interpretation of management conduct. Ethical adoption therefore requires clear policies concerning data access, model validation, bias assessment, documentation, and human oversight. Studies on AI in financial decision-making have highlighted the possibility that opaque technologies weaken ethical responsibility unless organizational controls are established [29]. Explainability can support accountability by making algorithmic decisions reviewable, but the existence of an explanation does not by itself ensure fairness or ethical compliance. Professional institutions and regulators must therefore define minimum expectations concerning transparency, reviewability, and the retention of auditor responsibility.

The maturity of the technological ecosystem in which audit firms operate also affects implementation. Explainable auditing requires access to high-quality data, reliable computing infrastructure, secure cloud services, suitable software, and specialized technology providers. Financial information systems must be sufficiently mature to produce complete, consistent, traceable, and machine-readable data. Data access and quality are particularly important because inaccurate or incomplete inputs may undermine both the model's predictions and the explanations derived from them [5]. The development of practical XAI maturity models reflects the need to evaluate readiness across technological, organizational, analytical, and evaluative dimensions rather than concentrating solely on model performance [15]. The feasibility of AI use in auditing is therefore partly determined by the broader infrastructure within audit firms, client organizations, and national regulatory systems.

Characteristics of the auditee can likewise influence the effectiveness of explainable auditing. Organizations with mature financial information systems, strong internal controls, standardized data structures, and cooperative management are more likely to provide the information required for accurate analysis. Conversely, fragmented systems, poor data governance, inconsistent records, and restricted access can reduce model performance and impair the quality of explanations. Because financial statement risk is closely related to the quality of internal processes and controls, the algorithmic assessment of risk should be interpreted in conjunction with the auditor's

understanding of the entity and its environment [4]. The model should therefore support, rather than replace, the auditor's examination of organizational context.

Professional and regulatory readiness further determines whether XAI-generated explanations can be relied upon in practice. Existing audit standards were developed primarily for human judgment and conventional evidence, and they may not explicitly address the validation, documentation, or review of artificial intelligence systems. Regulators and professional bodies must clarify the evidential status of algorithmic outputs, the documentation required for model-assisted judgments, and the extent to which XAI explanations can support audit conclusions. Research on accountability in audit firms underscores that responsibility must remain identifiable and reviewable throughout the engagement [30]. Future-oriented research in auditing similarly identifies technological change, regulation, professional competencies, and institutional adaptation as important forces shaping the profession [31]. Without an accepted professional framework, audit firms may adopt explainability practices inconsistently, reducing comparability and regulatory confidence.

The design of AI models is another fundamental component of explainable auditing. Model selection should reflect the audit objective, data characteristics, required level of transparency, cost of error, and professional consequences of a prediction. In some settings, intrinsically interpretable models may be preferable because their logic can be directly examined. In others, more complex models may be justified by higher predictive performance, provided that suitable post hoc explanations and validation procedures are available. The integration of artificial intelligence into audit processes should therefore involve a deliberate balance among accuracy, interpretability, stability, cost, and professional usability [2, 32]. An explainable model must also allow auditors to investigate alternative explanations, test contradictory evidence, and understand how changes in key inputs affect the resulting risk assessment.

When these conditions are satisfied, explainable auditing can generate important outcomes. It may increase the depth of risk analysis, improve the prioritization of audit procedures, reduce unnecessary testing, and optimize the allocation of personnel and time. Artificial intelligence-supported internal audit cycles can promote organizational learning and improve audit processes through iterative analysis and feedback [33]. XAI can also improve the consistency of professional judgment by making the factors underlying decisions more visible and comparable across engagements. In addition, it may support institutional knowledge transfer by documenting how risks were identified and why particular procedures were selected. These capabilities can improve audit quality while reducing cost and delay.

Nevertheless, the relationship between explainability and audit quality is unlikely to be direct or automatic. The benefits of XAI may depend on mediating factors such as auditor trust, the quality of professional judgment, and the quality of the explanations produced. A technically accurate explanation may have little value if it is not understandable to the auditor, and a clear explanation may still be misleading if it does not faithfully reflect the model. Similarly, auditor trust may improve the use of AI recommendations, but excessive trust may increase automation bias. Professional judgment may strengthen the interpretation of model outputs, but its effectiveness depends on the auditor's competence, skepticism, and willingness to revise an initial view. Accordingly, an integrated framework is needed to explain how antecedent conditions, core explainable-auditing components, mediating mechanisms, and positive and negative outcomes are related.

Although the existing literature addresses AI adoption, audit analytics, professional judgment, explainability techniques, regulatory transparency, and financial statement risk, these research streams remain fragmented. Some studies focus primarily on technical XAI methods, while others examine organizational adoption, ethical concerns,

or auditors' acceptance. Few studies integrate these dimensions into a comprehensive framework specifically designed for explainable auditing in financial statement risk assessment. Moreover, the available evidence does not sufficiently clarify how technological readiness, organizational capabilities, professional regulation, auditee characteristics, model design, explanation quality, auditor trust, and professional judgment interact to produce audit outcomes. A mixed-methods approach is therefore appropriate because qualitative synthesis can identify the conceptual dimensions of the phenomenon, while quantitative modeling can evaluate the relationships among the resulting constructs.

Accordingly, the aim of this study was to design and validate an explainable auditing framework for assessing financial statement risk based on explainable artificial intelligence through a sequential exploratory mixed-methods approach.

2. Methodology

This study is developmental–applied in terms of its objective because, while presenting a theoretical framework for conceptualizing explainable auditing for the assessment of financial statement risk based on explainable artificial intelligence, it also seeks to facilitate the framework's practical application to financial statements. In terms of its nature and overall research strategy, the study adopted a mixed-methods approach with a sequential exploratory design. Accordingly, qualitative data were first collected and analyzed, after which the quantitative phase was designed and implemented on the basis of the qualitative findings. This design enabled the researchers to use the richness of the qualitative data to develop a valid quantitative instrument and to integrate the findings through both inductive and deductive reasoning. Regarding the type of data, the study included both qualitative data, such as written materials and scientific documents, and quantitative data derived from questionnaire responses. The qualitative data were collected from the scientific literature, whereas the quantitative data were obtained through a survey. The qualitative phase employed a meta-synthesis approach to systematically review and synthesize previous studies. In the quantitative phase, partial least squares structural equation modeling (PLS-SEM) was used to examine the relationships among the components and validate the conceptual model. In terms of the time horizon, the study was cross-sectional. The qualitative data covered Persian-language articles published from 2021 onward and English-language articles published from 2020 onward. Regarding the degree of control over the variables, the study was descriptive and non-experimental because the data were collected without manipulating the independent variables, and the analyses were conducted on the basis of the existing relationships among the variables.

3. Findings and Results

Step 1: Formulation of the Research Questions

The substantive research questions were formulated around the dimensions of what, who, when, and how. These dimensions defined the substantive focus, study population, temporal boundaries, and analytical procedure of the meta-synthesis, as presented in Table 1.

Table 1. Research Questions and Scope of the Meta-Synthesis

Component	Research question or scope
What	What literature, models, frameworks, and techniques relating to explainable auditing, financial statement risk assessment, and explainable artificial intelligence have been used in the scientific literature? What frameworks have been used for the documentary content analysis of these subjects?

Who: Population under study	Citable scientific databases and scholarly information sources
International databases	Scopus, Web of Science, Google Scholar, Emerald, ScienceDirect, Taylor & Francis, and ProQuest
Persian-language databases	Scientific Information Database, Islamic World Science Citation Center, Comprehensive Portal of Humanities, Magiran, and Iranian Research Institute for Information Science and Technology
When	International studies published from 2020 onward and Persian-language studies published from 2021 onward
How	The data were analyzed through documentary analysis. Within the documentary content-analysis framework, the procedure involved extracting, coding, and qualitatively and quantitatively synthesizing data from the selected articles.

Step 2: Systematic Search for Research Sources

After the research questions had been specified, scientific sources and documents relevant to the research topic were systematically examined. International articles published from 2020 onward and Persian-language articles published from 2021 onward constituted the principal temporal scope of the search.

The search covered databases, journals, scholarly repositories, and search engines, including ResearchGate, ScienceDirect, ProQuest, Emerald, Google, and Google Scholar. International databases were searched extensively because no article specifically addressing an explainable auditing framework for assessing financial statement risk based on explainable artificial intelligence had been identified in Iranian academic journals.

Step 3: Selection of Articles and Sources Relevant to the Research

At this stage, the methodological quality of the retrieved sources was evaluated. The purpose was to exclude low-credibility sources and improve the precision, credibility, and analytical quality of the documentary content analysis.

The Critical Appraisal Skills Programme was used to evaluate the methodological quality of the primary studies. Its 10 appraisal domains concerned the clarity of the research objectives, appropriateness of the methodological rationale, research design, sampling method, data-collection procedure, researcher reflexivity, ethical considerations, rigor of data analysis, clarity of the findings, and overall value of the research.

Initially, 188 potentially relevant sources were identified. After the titles and thematic relevance of the sources had been assessed, 48 sources were excluded, leaving 140 sources. The abstracts of these 140 sources were then evaluated, resulting in the exclusion of 60 sources because they did not sufficiently correspond to the research topic or the domain of explainable auditing. The full texts of the remaining 80 articles were subsequently reviewed. Based on the quality-appraisal criteria, another 41 sources were excluded. Ultimately, 39 articles met the relevance and quality criteria and were selected for documentary content analysis.

According to the year-distribution data, the largest group of English-language articles was published in 2025, accounting for 11 articles and 28.21% of the total sample. Articles published in 2024 accounted for five articles and 12.82%. Among the Persian-language sources, three articles dated 2025 and three dated 2024 each accounted for 7.69%, while two Persian-language articles dated 2022 accounted for 5.13%. Two English-language articles published in 2022 and two published in 2021 each accounted for 5.13%. One article was reported for each of 2017, 2018, 2019, 2020, 2023, and 2026, with each accounting for 2.56%.

Step 4: Extraction of Concepts and Research-Related Codes

The information obtained from the selected articles and documents was classified according to the substantive focus, methodological procedure, temporal boundaries, and contextual scope of the research questions. Axial coding was then used to identify distinctive codes associated with explainable auditing for financial statement risk assessment based on explainable artificial intelligence.

An initial coding cycle yielded 73 distinctive codes. Through subsequent iterative review, recoding, and conceptual refinement, the comprehensive coding set was expanded to 102 extracted codes. The source-specific indicators underlying the coding process are reported in Table 2.

Table 2. Indicators Extracted from the Selected Studies

Researchers	Year	Extracted indicators
Sehat Barmcheh, Shamkhi, and Yalfani	2026	Step-by-step recording of audit procedures; association with the risk of material misstatement; greater analytical depth; transparency of decision logic; adjustment of judgment in response to new evidence
Morshedzadeh, Samiei, and Mashayekh	2023	Automated supporting evidence; reduction in audit risk
Adnan Hamoud, Piri, and Ashtab	2025	Reviewability of the analytical pathway; support for auditor decision-making
Begpanah, Esna-Ashari, Hoshi, and Asadi	2022	Reduction of complexity; reduced mechanical reliance on the system
Mansouri, Karami, and Yazdani	2024	Explanations adapted to the user's level; defensibility of decisions; reasoning in complex decisions
Nourallahzadeh and Sadeghi	2024	Use of professional auditing language; reduction of the expectation gap; consistency of judgment throughout the audit process
Vaezi and Akbari	2022	Focus on inherent and control risks; traceability of decisions; reconsideration of contradictory information
Abdi Baraftabi, Samadi, Shahverdiani, and Hajiha	2025	Strengthening human decision-making; reducing audit time and cost; cognitive flexibility
Riahinia, Daniali, and Haseli	2024	Documentation of judgment logic; optimization of resource allocation; compliance with professional standards
Bozorgasl, Pourmahdi, and Karimi Ghareh-Omar	2024	Decision trees; knowledge transfer within the audit firm; adherence to professional requirements
Anisykurlillah, Mukhibad, Jati, Rizkyana, Nurkhin, and Hapsoro	2025	Rule-based models; strengthening technological capabilities; alignment with best practices
Ariany	2025	SHAP and LIME; increasing complexity; alertness to fraud
Bagwe	2025	Local and global model explanations; illusion of understanding and misinterpretation; critical evaluation of evidence
Bracci, Tallaki, and Otia	2025	Focus on high-risk items; transfer or ambiguity of responsibility; professional skepticism
Carloni, Berti, and Colantonio	2025	Optimization of the audit plan; amplification or normalization of bias; comprehensibility of explanations
Dayeh	2025	Auditor–system question-and-answer functionality; accuracy of XAI model outputs; clarity of explanatory language
Edward	2025	Facilitation of audit-team communication; reliability of model analyses; understanding the reasons underlying model outputs
Eulerich, Pawlowski, Waddoups, and Wood	2022	Reduction of ambiguity; stability of model performance; relevance to audit risks
Golubovskij	2024	Algorithmic accountability; comprehensibility of model logic; usefulness for decision-making
Kalyanathaya and Prasad	2022	Identification of data bias; transparency of algorithms and processes; focus on key factors
Kassar and Jizi	2025	Alignment with professional ethical principles; usefulness of the model in auditing; explanation of the reasons for results
Kesari, Sele, Ash, and Bechtold	2024	Clean and structured data; increased audit efficiency; consistency with audit findings
Kokina and Davenport	2017	Data traceability; assistance in risk identification; consistency with professional standards
Langer, Baum, Hartmann, Hessel, Speith, and Wahl	2024	Ethical and privacy considerations; predictability of model behavior; guidance for subsequent actions
Liu, Liu, Lu, and Olofsson	2025	Availability of valid explanation techniques; estimability of model results; assistance in designing audit tests
Lyu and Wu	2025	Explanation-validation criteria; accuracy of professional judgment; improvement of the audit process

Miller	2019	Local and global model explanations; access to cloud infrastructure; explanation of the effects of variables on outputs
Mumuni and Mumuni	2025	Auditors familiar with data concepts; existence of reputable technology companies; identification of influential variables
Muñoz-Ordóñez, Cobos, Vidal-Rojas, and Herrera	2025	Data specialists familiar with auditing principles; transparency of model-processing procedures; availability of specialist personnel
Ramos and Esther	2024	Culture of learning and innovation; applied training courses
Razali, Sulaiman, Abdul Manan, and Said	2025	Senior-management support; traceability and rich metadata
Rudin, Chen, Chen, Huang, Semenova, and Zhong	2021	Updating auditing standards
Salih, Raisi-Estabragh, Boscolo Galazzo, Radeva, Petersen, Lekadir, and Menegaz	2024	Practical guidance from professional institutions
Schumann and Marx Gómez	2021	Clarity regarding the auditor's ultimate responsibility
Taofeek	2025	Recognition of explanations by regulatory authorities
Thành, Duyen, and Hang	2025	Maturity of financial information systems
Whitaker, Brooks, Price, Reynolds, and Paul	2020	Effectiveness of internal controls
Zavitsanos, Spyropoulou, Giannakopoulos, and Paliouras	2025	Access to data
Zhong and Goel	2024	Shared understanding of audit objectives

Step 5: Analysis and Synthesis of the Qualitative Findings

The contents of the selected articles and sources were examined in detail, and codes associated with the research keywords were identified. The extracted factors were initially treated as independent codes and were subsequently grouped according to their semantic and conceptual similarities.

The final analytical process produced 102 extracted codes. These codes were consolidated into 42 initial codes, which were then classified into 18 organizing codes. The organizing codes were explainability and interpretability of results; professional justifiability and alignment with audit logic; selection and design of artificial intelligence models; applicability in the audit process; trust-building and professional acceptance; organizational and infrastructural requirements; technical and infrastructural readiness; organizational and human readiness; professional and regulatory readiness; auditee characteristics; national technology-ecosystem maturity; audit-quality outcomes; professional credibility and accountability outcomes; operational and economic outcomes; negative outcomes and risks; auditor trust in XAI models; auditor professional-judgment quality; and XAI explanation quality.

The 18 organizing codes were ultimately classified under four overarching codes: explainable auditing based on explainable artificial intelligence, antecedents, outcomes, and mediating factors. The final dimensions, categories, initial codes, and extracted concepts are shown in Table 3.

Table 3. Final Dimensions, Organizing Codes, Initial Codes, and Extracted Concepts

Overarching code	Organizing code	Initial codes	Extracted concepts
Explainable auditing based on explainable artificial intelligence	Explainability and interpretability of results	Audit-process traceability	Step-by-step recording of audit procedures; automated supporting evidence; reviewability of the analytical pathway
		Comprehensible formulation of explanations	Reduction of complexity; explanation adapted to the user's level; use of professional auditing language
	Professional justifiability and alignment with audit logic	Linking explanations to audit risks	Association with the risk of material misstatement; focus on inherent and control risks

		Support for auditors' professional judgment	Strengthening human decision-making; documentation of judgment logic
	Selection and design of artificial intelligence models	Intrinsically interpretable models	Decision trees; rule-based models
		Complex models with post hoc explanations	SHAP and LIME; local and global model explanations
	Applicability in the audit process	Guidance and prioritization of audit tests	Focus on high-risk items; optimization of the audit plan
		Human–system interaction	Auditor–system question-and-answer functionality; facilitation of audit-team communication
	Trust-building and professional acceptance	Perceived system credibility	Reduction of ambiguity; algorithmic accountability
		Fairness and bias control	Identification of data bias; alignment with professional ethical principles
	Organizational and infrastructural requirements	Data quality and governance	Clean and structured data; data traceability
		Organizational and ethical capabilities	Auditor competencies; ethical and privacy considerations
Antecedents	Technical and infrastructural readiness	Explainability capability of AI models	Availability of valid explanation techniques; explanation-validation criteria
		Integrated, high-quality data infrastructure	Clean and structured data; traceability and rich metadata
	Organizational and human readiness	Teams' hybrid competencies	Auditors familiar with data concepts; data specialists familiar with auditing principles
		Supportive organizational culture	Culture of learning and innovation; senior-management support
	Professional and regulatory readiness	Professional standards and guidance	Updating auditing standards; practical guidance from professional institutions
		Receptive legal and regulatory framework	Clarity regarding the auditor's ultimate responsibility; recognition of explanations by regulators
	Auditee characteristics	Maturity of the client's information systems	Maturity of financial information systems; effectiveness of internal controls
		Cooperation of auditee management	Access to data; shared understanding of audit objectives
	National technology-ecosystem maturity	Technical infrastructure and services	Access to cloud infrastructure; existence of reputable technology companies
		Human capital and technical knowledge	Availability of specialist personnel; applied training courses
Outcomes	Audit-quality outcomes	Improved analysis and risk assessment	Greater analytical depth; reduction in audit risk
		Strengthened professional judgment	Support for auditor decision-making; reduced mechanical reliance on the system
	Professional credibility and accountability outcomes	Greater transparency and public trust	Defensibility of decisions; reduction of the expectation gap
		Accountability and responsibility	Transparency of decision logic; traceability of decisions
	Operational and economic outcomes	Increased audit efficiency	Reduction of audit time and cost; optimization of resource allocation
		Organizational learning and development	Knowledge transfer within the audit firm; strengthening technological capabilities
	Negative outcomes and risks	Technical and cognitive challenges	Increasing complexity; illusion of understanding and misinterpretation

		Professional and ethical challenges	Transfer or ambiguity of responsibility; amplification or normalization of bias
Mediating factors	Auditor trust in XAI models	Model reliability and transparency	Accuracy of XAI model outputs; reliability of model analyses; stability of model performance; comprehensibility of model logic; transparency of algorithms and processes
		Perceived benefit and controllability	Usefulness of the model in auditing; increased audit efficiency; assistance in risk identification; preservation of auditor control over the process
		Predictability of model behavior	Predictability of model behavior; estimability of model results
	Auditor professional-judgment quality	Decision accuracy	Accuracy of professional judgment; reasoning in complex decisions; consistency of judgment throughout the audit process
		Adaptability and flexibility	Adjustment of judgment in response to new evidence; reconsideration of contradictory information; cognitive flexibility
		Institutionalization of standards	Compliance with professional standards; adherence to professional requirements; alignment with best practices
		Professional skepticism	Alertness to fraud; critical evaluation of evidence; professional skepticism
	XAI explanation quality	Clarity and comprehensibility	Comprehensibility of explanations; clarity of explanatory language; understanding the reasons underlying model outputs
		Relevance and appropriateness	Relevance to audit risks; usefulness for decision-making; focus on key factors; explanation of the reasons for results
		Logical consistency	Consistency with audit findings; consistency with professional standards; logical coherence of explanations
		Actionability	Guidance for subsequent actions; assistance in designing audit tests; improvement of the audit process
		Degree of algorithmic transparency	Explanation of the effects of variables on outputs; identification of influential variables; transparency of model-processing procedures

Step 6: Quality Control

Although the selected sources and extracted codes were continuously reviewed and recoded to ensure the quality of the qualitative findings, Cohen's kappa coefficient was also used to assess intercoder reliability and internal consistency.

The extracted codes and distinctive categories were submitted to two experts. After their assessments had been collected, Cohen's kappa was calculated based on their agreements and disagreements regarding the codes and categories. The resulting Cohen's kappa coefficient was 0.742, with $p < .001$. Because this coefficient exceeded 0.60, it indicated substantial and acceptable agreement between the experts.

Step 7: Presentation of the Meta-Synthesis Findings

The findings of the meta-synthesis were structured around four overarching codes. The first overarching code, explainable auditing based on explainable artificial intelligence, emphasized the explainability and interpretability of results, professional justifiability and alignment with audit logic, selection and design of AI models, applicability in the audit process, trust-building and professional acceptance, and organizational and infrastructural requirements.

The practical objective of this dimension is to ensure that XAI systems can be effectively incorporated into audit planning, risk assessment, test prioritization, and human–system interaction while preserving the auditor's professional responsibility.

The second overarching code, antecedents, encompassed the conditions required for the successful implementation of explainable auditing. These conditions included technical and infrastructural readiness, organizational and human readiness, professional and regulatory readiness, auditee characteristics, and the maturity of the national technology ecosystem.

The third overarching code, outcomes, indicated that successful implementation of explainable auditing could improve audit quality, professional credibility and accountability, audit efficiency, and organizational learning and development while reducing audit time and cost. Nevertheless, technical and cognitive risks, as well as professional and ethical challenges, may also arise and require active governance.

The fourth overarching code, mediating factors, comprised auditor trust in XAI models, auditor professional-judgment quality, and XAI explanation quality. These factors may determine the extent to which explainable auditing capabilities translate into desirable professional, operational, and audit-quality outcomes.

Overall, the qualitative findings provide an integrated theoretical and practical framework for enhancing audit quality and efficiency through the responsible integration of explainable artificial intelligence.

The demographic characteristics examined in the quantitative phase included gender, age, marital status, educational attainment, and professional experience. These characteristics provide information about the composition, diversity, and representativeness of the sample of 98 auditors, accountants, and information technology specialists.

The percentages in Table 4 were recalculated directly from the reported frequencies and the sample size of 98 to correct minor inconsistencies in the supplied percentages.

Table 4. Demographic Characteristics of the Respondents

Variable	Category	*n*	%
Gender	Women	56	57.1
	Men	42	42.9
Age	Younger than 30 years	22	22.4
	30–39 years	44	44.9
	40–49 years	23	23.5
	50 years or older	9	9.2
Marital status	Married	66	67.3
	Single	32	32.7
Educational attainment	Bachelor's degree	16	16.3
	Master's degree	63	64.3
	Doctoral degree or higher	19	19.4
Professional experience	Fewer than 5 years	25	25.5
	5–10 years	36	36.7
	More than 10 years	37	37.8

Of the 98 respondents, 56 were women and 42 were men. The largest age group consisted of respondents aged 30–39 years, followed by those aged 40–49 years, those younger than 30 years, and those aged 50 years or older.

Most participants were married, and nearly two-thirds held a master's degree. Regarding professional experience, 37 participants had more than 10 years of experience, 36 had between 5 and 10 years of experience, and 25 had fewer than 5 years of experience.

PLS-SEM model assessment was conducted in three stages: assessment of the measurement model, assessment of the structural model, and assessment of the model's overall predictive performance.

Factor loadings represent the magnitude and direction of the relationships between observed indicators and their corresponding latent constructs. Loadings of 0.40 or higher were considered minimally acceptable. All indicator loadings exceeded 0.40, indicating acceptable indicator reliability.

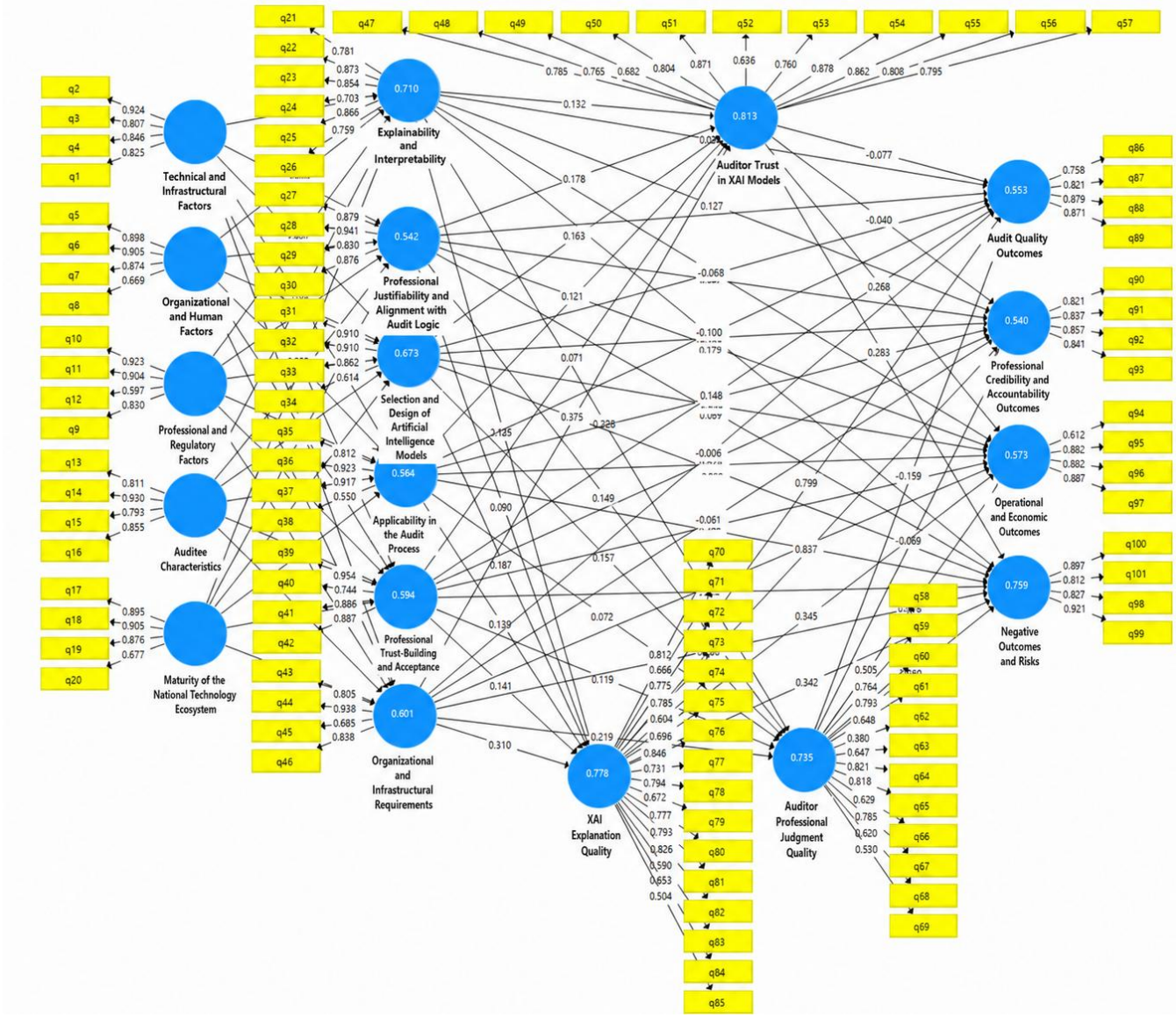


Figure 1. Measurement model in the coefficient-significance estimation mode.

Cronbach's alpha was used to assess the internal consistency of the indicators associated with each construct. Values greater than 0.70 were interpreted as evidence of acceptable reliability.

Composite reliability was also used to evaluate internal consistency while accounting for differences in indicator loadings. Values exceeding 0.70 indicated satisfactory construct reliability.

The average variance extracted represented the proportion of indicator variance explained by the corresponding latent construct. An AVE value of at least 0.50 indicated that a construct explained more than half of the variance in its indicators and therefore demonstrated convergent validity.

The rho_A coefficient was additionally evaluated as a reliability measure positioned between Cronbach's alpha and composite reliability. Values greater than 0.70 were considered acceptable.

Table 5. Reliability and Convergent-Validity Indices for the Research Constructs

Construct	Cronbach's α	rho_A	Composite reliability	AVE
Auditor trust in XAI models	0.938	0.944	0.947	0.623
Professional trust-building and acceptance	0.893	0.924	0.926	0.759
Organizational and infrastructural requirements	0.837	0.872	0.891	0.674
AI model selection and design	0.849	0.895	0.899	0.694
National technology-ecosystem maturity	0.862	0.890	0.907	0.712
Professional justifiability and alignment with audit logic	0.908	0.946	0.933	0.778
Explainability and interpretability	0.895	0.910	0.919	0.655
Professional and regulatory factors	0.833	0.864	0.892	0.679
Organizational and human factors	0.860	0.891	0.906	0.709
Technical and infrastructural factors	0.880	0.922	0.913	0.726
Auditee characteristics	0.879	0.934	0.911	0.721
Professional credibility and accountability outcomes	0.860	0.860	0.905	0.704
Operational and economic outcomes	0.834	0.855	0.892	0.679
Audit-quality outcomes	0.854	0.867	0.901	0.695
Negative outcomes and risks	0.887	0.890	0.923	0.749
Applicability in the audit process	0.817	0.853	0.884	0.663
XAI explanation quality	0.938	0.944	0.946	0.528
Auditor professional-judgment quality	0.908	0.919	0.924	0.509

As shown in Table 5, Cronbach's alpha exceeded 0.70 for every construct, indicating satisfactory internal consistency. Composite-reliability values also exceeded 0.70 for all constructs, confirming construct reliability.

All AVE values were greater than 0.50, indicating that each latent construct explained more than half of the variance in its associated indicators. Convergent validity was therefore supported. All rho_A coefficients were also above 0.70, providing additional evidence of internal-consistency reliability.

The Fornell–Larcker criterion assesses discriminant validity by comparing the square root of each construct's AVE with its correlations with other constructs. Discriminant validity is supported when the square root of AVE is greater than all interconstruct correlations involving that construct.

The heterotrait–monotrait ratio of correlations was also examined. HTMT values below 0.90 are generally considered acceptable, although a stricter threshold of 0.85 is sometimes recommended (Henseler et al., 2015).

Because the complete 18 × 18 Fornell–Larcker and HTMT matrices would be excessively large, Table 6 reports the square root of AVE, the highest interconstruct correlation, and the highest HTMT value for each construct.

Table 6. Summary of Fornell–Larcker and HTMT Discriminant-Validity Results

Construct	√AVE	Highest correlation	Construct associated with highest correlation	Highest HTMT
Auditor trust in XAI models	0.789	0.851	XAI explanation quality	0.896
Professional trust-building and acceptance	0.871	0.755	XAI explanation quality	0.882
Organizational and infrastructural requirements	0.821	0.829	Auditor trust in XAI models	0.870
AI model selection and design	0.833	0.815	Auditor trust in XAI models	0.894
National technology-ecosystem maturity	0.844	0.818	Explainability and interpretability	0.847
Professional justifiability and alignment with audit logic	0.882	0.722	AI model selection and design	0.764
Explainability and interpretability	0.809	0.818	National technology-ecosystem maturity	0.831
Professional and regulatory factors	0.824	0.823	Organizational and human factors	0.860
Organizational and human factors	0.842	0.823	Professional and regulatory factors	0.861
Technical and infrastructural factors	0.852	0.695	Organizational and human factors	0.750

Auditee characteristics	0.849	0.705	Auditor professional-judgment quality	0.750
Professional credibility and accountability outcomes	0.839	0.784	Audit-quality outcomes	0.810
Operational and economic outcomes	0.824	0.822	Negative outcomes and risks	0.844
Audit-quality outcomes	0.834	0.784	Professional credibility and accountability outcomes	0.812
Negative outcomes and risks	0.865	0.826	XAI explanation quality	0.897
Applicability in the audit process	0.815	0.769	AI model selection and design	0.882
XAI explanation quality	0.727	0.851	Auditor trust in XAI models	0.897
Auditor professional-judgment quality	0.713	0.824	XAI explanation quality	0.880

The Fornell–Larcker criterion was fully satisfied for several constructs, including professional trust-building and acceptance, AI model selection and design, national technology-ecosystem maturity, professional justifiability and alignment with audit logic, professional and regulatory factors, organizational and human factors, technical and infrastructural factors, auditee characteristics, professional credibility and accountability outcomes, audit-quality outcomes, negative outcomes and risks, and applicability in the audit process.

However, it was not satisfied for auditor trust in XAI models, organizational and infrastructural requirements, explainability and interpretability, XAI explanation quality, or auditor professional-judgment quality because their highest correlations exceeded their respective square roots of AVE. The operational and economic outcomes construct also demonstrated only a marginal distinction because its square root of AVE was 0.824 and its correlation with negative outcomes and risks was 0.822.

All reported HTMT values remained below the 0.90 threshold, with the highest value being 0.897. Therefore, discriminant validity was broadly supported under the 0.90 criterion. Nevertheless, several values exceeded the stricter 0.85 threshold, particularly those involving auditor trust, XAI explanation quality, negative outcomes and risks, applicability in the audit process, AI model selection and design, and auditor professional-judgment quality. These results indicate substantial conceptual proximity among several constructs and suggest that their empirical distinctiveness should be interpreted cautiously.

The structural model was assessed using the coefficient of determination, Stone–Geisser's predictive-relevance coefficient, and Cohen's effect size.

The coefficient of determination indicates the proportion of variance in an endogenous construct explained by its predictor constructs. Values of approximately 0.19, 0.33, and 0.67 were interpreted as weak, moderate, and strong, respectively (Davari & Rezazadeh, 2014).

Stone–Geisser's Q^2 coefficient was used to assess the model's out-of-sample predictive relevance. Values of approximately 0.02, 0.15, and 0.35 indicate weak, moderate, and strong predictive relevance, respectively.

Table 7. Explanatory Power and Predictive Relevance of the Structural Model

Endogenous construct	R^2	Adjusted R^2	Explanatory power	Q^2	Predictive relevance
Auditor trust in XAI models	0.813	0.810	Strong	0.470	Strong
Professional trust-building and acceptance	0.594	0.588	Moderate	0.416	Strong
Organizational and infrastructural requirements	0.601	0.596	Moderate	0.368	Strong
AI model selection and design	0.673	0.669	Strong	0.435	Strong
Professional justifiability and alignment with audit logic	0.542	0.535	Moderate	0.381	Strong
Explainability and interpretability	0.710	0.706	Strong	0.425	Strong
Professional credibility and accountability outcomes	0.540	0.529	Moderate	0.352	Strong
Operational and economic outcomes	0.573	0.563	Moderate	0.351	Strong
Audit-quality outcomes	0.553	0.542	Moderate	0.353	Strong

Negative outcomes and risks	0.759	0.753	Strong	0.532	Strong
Applicability in the audit process	0.564	0.558	Moderate	0.351	Strong
XAI explanation quality	0.778	0.774	Strong	0.381	Strong
Auditor professional-judgment quality	0.735	0.731	Strong	0.346	Moderate

National technology-ecosystem maturity, professional and regulatory factors, organizational and human factors, technical and infrastructural factors, and auditee characteristics were modeled as exogenous constructs. Consequently, R^2 and Q^2 values were not calculated for these variables.

The R^2 results indicate that the model had strong explanatory power for auditor trust in XAI models, AI model selection and design, explainability and interpretability, negative outcomes and risks, XAI explanation quality, and auditor professional-judgment quality. The remaining endogenous constructs demonstrated moderate explanatory power.

The highest R^2 value was obtained for auditor trust in XAI models, indicating that 81.3% of its variance was explained by the predictor constructs. XAI explanation quality and negative outcomes and risks also had high R^2 values of 0.778 and 0.759, respectively.

Almost all Q^2 values exceeded 0.35, indicating strong predictive relevance. The Q^2 value for auditor professional-judgment quality was 0.346 and was therefore marginally below the 0.35 threshold but remained well above the criterion for moderate predictive relevance. Overall, the model demonstrated substantial explanatory and predictive capability.

Cohen's f^2 coefficient was used to assess the change in an endogenous construct's R^2 resulting from the inclusion of a particular predictor. Values of 0.02, 0.15, and 0.35 represent small, medium, and large effects, respectively.

Most individual effect sizes were small or negligible, which is expected in a complex model containing numerous simultaneous predictors. Nevertheless, several theoretically important effects were identified.

National technology-ecosystem maturity had a large effect on explainability and interpretability, $f^2 = 0.407$. It also had medium effects on professional trust-building and acceptance, $f^2 = 0.216$, and effects approaching the medium threshold on professional justifiability and alignment with audit logic, $f^2 = 0.147$.

Organizational and infrastructural requirements had a medium effect on auditor trust in XAI models, $f^2 = 0.259$, and a medium effect on XAI explanation quality, $f^2 = 0.150$.

XAI explanation quality had medium effects on professional credibility and accountability outcomes, $f^2 = 0.289$, and audit-quality outcomes, $f^2 = 0.270$. Its effects on negative outcomes and risks, $f^2 = 0.092$, and operational and economic outcomes, $f^2 = 0.053$, were small.

National technology-ecosystem maturity also had small effects on applicability in the audit process, $f^2 = 0.104$, AI model selection and design, $f^2 = 0.067$, and organizational and infrastructural requirements, $f^2 = 0.052$.

Other notable small effects included explainability and interpretability on auditor professional-judgment quality, $f^2 = 0.092$; technical and infrastructural factors on AI model selection and design, $f^2 = 0.082$; professional justifiability and alignment with audit logic on auditor trust, $f^2 = 0.076$; organizational and infrastructural requirements on auditor professional-judgment quality, $f^2 = 0.063$; and auditor trust in XAI models on negative outcomes and risks, $f^2 = 0.056$.

Accordingly, although many individual pathways had weak effect sizes, several medium and large effects confirmed the substantive importance of technology-ecosystem maturity, organizational and infrastructural requirements, and XAI explanation quality.

The structural model was estimated using bootstrapping. The significance of each path was evaluated using the t statistic and corresponding p value. Paths with an absolute t value greater than 1.96 and $p < .05$ were considered statistically significant.

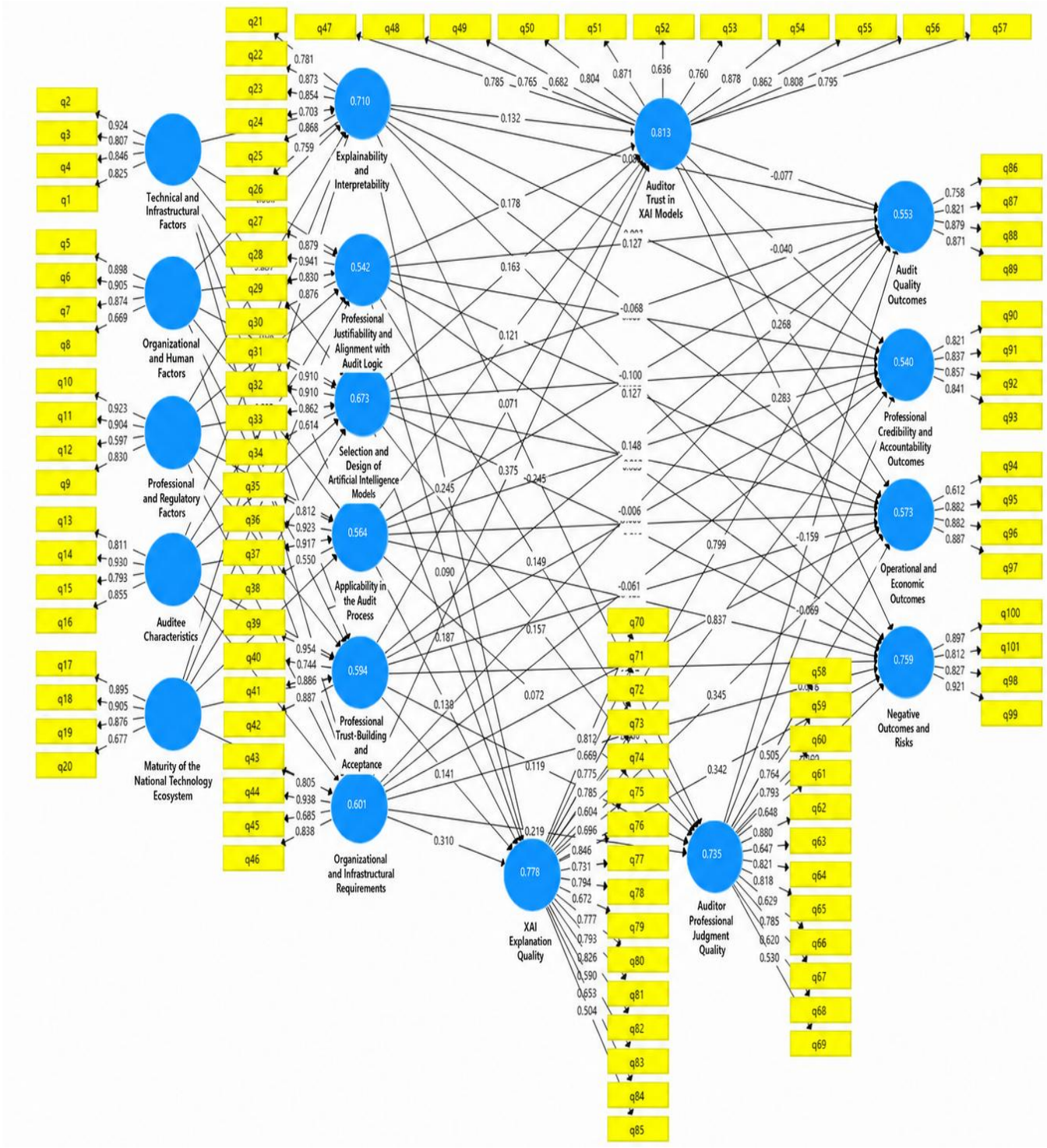


Figure 2. Structural model in the standard-coefficient mode.

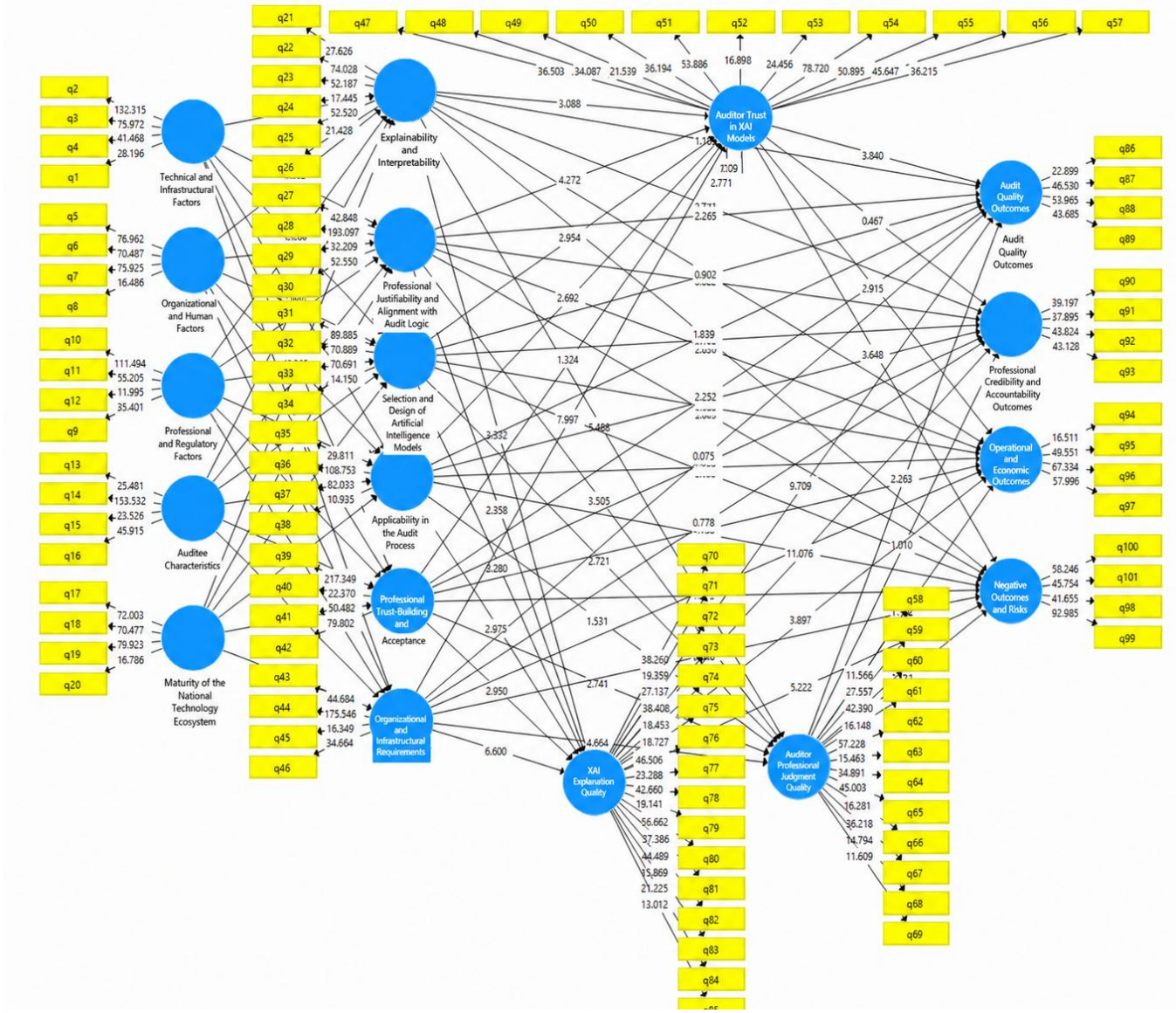


Figure 3. Structural model in the bootstrap significance-estimation mode.

The direct-path results are presented in Table 8.

Table 8. Results of the Structural-Model Path Tests

Structural path	β	SD	*t*	*p*	Result
Auditor trust in XAI models → Professional credibility and accountability outcomes	−0.040	0.085	0.467	.641	Not significant
Auditor trust in XAI models → Operational and economic outcomes	0.268	0.092	2.915	.004	Significant
Auditor trust in XAI models → Audit-quality outcomes	−0.077	0.092	0.840	.401	Not significant
Auditor trust in XAI models → Negative outcomes and risks	0.283	0.078	3.648	< .001	Significant
Professional trust-building and acceptance → Auditor trust in XAI models	0.071	0.053	1.324	.186	Not significant
Professional trust-building and acceptance → Professional credibility and accountability outcomes	0.923	0.008	111.494	< .001	Significant
Professional trust-building and acceptance → Operational and economic outcomes	0.897	0.015	58.248	< .001	Significant
Professional trust-building and acceptance → Audit-quality outcomes	0.148	0.066	2.252	.025	Significant
Professional trust-building and acceptance → Negative outcomes and risks	0.102	0.054	1.898	.058	Not significant
Professional trust-building and acceptance → XAI explanation quality	0.141	0.048	2.950	.003	Significant

Professional trust-building and acceptance → Auditor professional-judgment quality	0.119	0.043	2.741	.006	Significant
Organizational and infrastructural requirements → Auditor trust in XAI models	0.375	0.047	7.997	< .001	Significant
Organizational and infrastructural requirements → Professional credibility and accountability outcomes	−0.061	0.078	0.778	.437	Not significant
Organizational and infrastructural requirements → Operational and economic outcomes	0.854	0.016	52.187	< .001	Significant
Organizational and infrastructural requirements → Audit-quality outcomes	−0.006	0.080	0.075	.940	Not significant
Organizational and infrastructural requirements → Negative outcomes and risks	0.703	0.040	17.445	< .001	Significant
Organizational and infrastructural requirements → XAI explanation quality	0.310	0.047	6.600	< .001	Significant
Organizational and infrastructural requirements → Auditor professional-judgment quality	0.219	0.047	4.664	< .001	Significant
AI model selection and design → Auditor trust in XAI models	0.163	0.055	2.954	.003	Significant
AI model selection and design → Professional credibility and accountability outcomes	0.905	0.013	70.477	< .001	Significant
AI model selection and design → Operational and economic outcomes	0.878	0.011	79.923	< .001	Significant
AI model selection and design → Audit-quality outcomes	−0.068	0.076	0.902	.367	Not significant
AI model selection and design → Negative outcomes and risks	−0.049	0.057	0.868	.386	Not significant
AI model selection and design → XAI explanation quality	0.187	0.057	3.280	.001	Significant
AI model selection and design → Auditor professional-judgment quality	0.157	0.058	2.721	.007	Significant
National technology-ecosystem maturity → Professional trust-building and acceptance	0.500	0.059	8.481	< .001	Significant
National technology-ecosystem maturity → Organizational and infrastructural requirements	0.242	0.065	3.725	< .001	Significant
National technology-ecosystem maturity → AI model selection and design	0.250	0.055	4.547	< .001	Significant
National technology-ecosystem maturity → Professional justifiability and alignment with audit logic	0.439	0.061	7.236	< .001	Significant
National technology-ecosystem maturity → Explainability and interpretability	0.580	0.058	10.066	< .001	Significant
National technology-ecosystem maturity → Applicability in the audit process	0.360	0.061	5.862	< .001	Significant
Professional justifiability and alignment with audit logic → Auditor trust in XAI models	0.178	0.042	4.272	< .001	Significant
Professional justifiability and alignment with audit logic → Professional credibility and accountability outcomes	0.047	0.057	0.822	.412	Not significant
Professional justifiability and alignment with audit logic → Operational and economic outcomes	0.154	0.058	2.636	.009	Significant
Professional justifiability and alignment with audit logic → Audit-quality outcomes	0.127	0.056	2.265	.024	Significant
Professional justifiability and alignment with audit logic → Negative outcomes and risks	0.155	0.042	3.649	< .001	Significant
Professional justifiability and alignment with audit logic → XAI explanation quality	0.090	0.038	2.358	.019	Significant
Professional justifiability and alignment with audit logic → Auditor professional-judgment quality	0.149	0.042	3.505	< .001	Significant
Explainability and interpretability → Auditor trust in XAI models	0.132	0.043	3.088	.002	Significant
Explainability and interpretability → Professional credibility and accountability outcomes	0.812	0.018	45.754	< .001	Significant
Explainability and interpretability → Operational and economic outcomes	0.904	0.014	65.205	< .001	Significant

Explainability and interpretability → Audit-quality outcomes	0.597	0.050	11.995	< .001	Significant
Explainability and interpretability → Negative outcomes and risks	0.811	0.032	25.481	< .001	Significant
Explainability and interpretability → XAI explanation quality	0.145	0.043	3.332	.001	Significant
Explainability and interpretability → Auditor professional-judgment quality	0.274	0.053	5.138	< .001	Significant
Professional and regulatory factors → Professional trust-building and acceptance	0.251	0.063	3.992	< .001	Significant
Professional and regulatory factors → Organizational and infrastructural requirements	0.152	0.071	2.132	.033	Significant
Professional and regulatory factors → AI model selection and design	0.139	0.058	2.402	.017	Significant
Professional and regulatory factors → Professional justifiability and alignment with audit logic	0.187	0.067	2.791	.005	Significant
Professional and regulatory factors → Explainability and interpretability	0.207	0.051	4.034	< .001	Significant
Professional and regulatory factors → Applicability in the audit process	0.930	0.006	153.532	< .001	Significant
Organizational and human factors → Professional trust-building and acceptance	0.793	0.034	23.526	< .001	Significant
Organizational and human factors → Organizational and infrastructural requirements	0.191	0.069	2.767	.006	Significant
Organizational and human factors → AI model selection and design	0.210	0.058	3.593	< .001	Significant
Organizational and human factors → Professional justifiability and alignment with audit logic	−0.137	0.063	2.160	.031	Significant negative effect
Organizational and human factors → Explainability and interpretability	0.000	0.046	0.009	.993	Not significant
Organizational and human factors → Applicability in the audit process	0.302	0.063	4.773	< .001	Significant
Technical and infrastructural factors → Professional trust-building and acceptance	0.148	0.055	2.693	.007	Significant
Technical and infrastructural factors → Organizational and infrastructural requirements	0.142	0.058	2.471	.014	Significant
Technical and infrastructural factors → AI model selection and design	0.231	0.048	4.772	< .001	Significant
Technical and infrastructural factors → Professional justifiability and alignment with audit logic	0.200	0.058	3.462	.001	Significant
Technical and infrastructural factors → Explainability and interpretability	0.100	0.047	2.131	.034	Significant
Technical and infrastructural factors → Applicability in the audit process	0.191	0.054	3.506	< .001	Significant
Auditee characteristics → Professional trust-building and acceptance	−0.051	0.057	0.896	.371	Not significant
Auditee characteristics → Organizational and infrastructural requirements	0.178	0.058	3.074	.002	Significant
Auditee characteristics → AI model selection and design	0.130	0.046	2.837	.005	Significant
Auditee characteristics → Professional justifiability and alignment with audit logic	0.156	0.057	2.750	.006	Significant
Auditee characteristics → Explainability and interpretability	0.855	0.019	45.915	< .001	Significant
Auditee characteristics → Applicability in the audit process	0.895	0.012	72.003	< .001	Significant
Applicability in the audit process → Auditor trust in XAI models	0.121	0.045	2.692	.007	Significant
Applicability in the audit process → Professional credibility and accountability outcomes	−0.195	0.065	2.987	.003	Significant negative effect
Applicability in the audit process → Operational and economic outcomes	0.677	0.040	16.786	< .001	Significant
Applicability in the audit process → Audit-quality outcomes	−0.100	0.054	1.839	.067	Not significant

Applicability in the audit process → Negative outcomes and risks	0.924	0.007	132.315	< .001	Significant
Applicability in the audit process → XAI explanation quality	0.139	0.047	2.975	.003	Significant
Applicability in the audit process → Auditor professional-judgment quality	0.072	0.047	1.531	.127	Not significant
XAI explanation quality → Professional credibility and accountability outcomes	0.837	0.076	11.076	< .001	Significant
XAI explanation quality → Operational and economic outcomes	0.345	0.089	3.897	< .001	Significant
XAI explanation quality → Audit-quality outcomes	0.799	0.082	9.709	< .001	Significant
XAI explanation quality → Negative outcomes and risks	0.342	0.066	5.222	< .001	Significant
Auditor professional-judgment quality → Professional credibility and accountability outcomes	−0.069	0.068	1.010	.313	Not significant
Auditor professional-judgment quality → Operational and economic outcomes	0.781	0.028	27.626	< .001	Significant
Auditor professional-judgment quality → Audit-quality outcomes	−0.159	0.070	2.263	.024	Significant negative effect
Auditor professional-judgment quality → Negative outcomes and risks	0.873	0.012	74.028	< .001	Significant

The results indicate that a substantial proportion of the proposed relationships were statistically significant. Explainability and interpretability, XAI explanation quality, organizational and infrastructural requirements, and AI model selection and design emerged as particularly influential constructs.

Explainability and interpretability had significant positive effects on professional credibility and accountability outcomes, $\beta = 0.812$, operational and economic outcomes, $\beta = 0.904$, audit-quality outcomes, $\beta = 0.597$, and negative outcomes and risks, $\beta = 0.811$. It also significantly enhanced auditor trust in XAI models, XAI explanation quality, and auditor professional-judgment quality.

XAI explanation quality had significant positive effects on all four outcome categories. Its strongest effect concerned professional credibility and accountability outcomes, $\beta = 0.837$, followed by audit-quality outcomes, $\beta = 0.799$.

Professional trust-building and acceptance had very strong positive effects on professional credibility and accountability outcomes, $\beta = 0.923$, and operational and economic outcomes, $\beta = 0.897$. It also had smaller but significant positive effects on audit-quality outcomes, XAI explanation quality, and auditor professional-judgment quality.

Organizational and infrastructural requirements significantly increased auditor trust in XAI models, operational and economic outcomes, negative outcomes and risks, XAI explanation quality, and auditor professional-judgment quality. However, their direct effects on professional credibility and accountability outcomes and audit-quality outcomes were not significant.

AI model selection and design significantly influenced auditor trust, professional credibility and accountability outcomes, operational and economic outcomes, XAI explanation quality, and professional-judgment quality. Its direct effects on audit-quality outcomes and negative outcomes and risks were not statistically significant.

National technology-ecosystem maturity had significant positive effects on all six endogenous implementation dimensions examined: professional trust-building and acceptance, organizational and infrastructural requirements, AI model selection and design, professional justifiability and alignment with audit logic, explainability and

interpretability, and applicability in the audit process. This finding identifies ecosystem maturity as an important upstream condition for the implementation of explainable auditing.

Professional and regulatory factors also had significant positive effects on all implementation dimensions. Their strongest effect was on applicability in the audit process, $\beta = 0.930$.

Organizational and human factors significantly improved professional trust-building and acceptance, organizational and infrastructural requirements, AI model selection and design, and applicability in the audit process. Their relationship with professional justifiability and alignment with audit logic was significant but negative, $\beta = -0.137$, while their direct effect on explainability and interpretability was not significant.

Technical and infrastructural factors had significant positive effects on all six implementation dimensions. Auditee characteristics also significantly influenced organizational and infrastructural requirements, AI model selection and design, professional justifiability, explainability and interpretability, and applicability in the audit process. However, auditee characteristics did not significantly affect professional trust-building and acceptance.

Applicability in the audit process increased auditor trust, operational and economic outcomes, negative outcomes and risks, and XAI explanation quality. Its effect on professional credibility and accountability outcomes was significant but negative, $\beta = -0.195$. Its effects on audit-quality outcomes and auditor professional-judgment quality were not significant.

Auditor trust in XAI models significantly improved operational and economic outcomes but did not significantly affect professional credibility and accountability outcomes or audit-quality outcomes. It was also positively associated with negative outcomes and risks, suggesting that greater reliance on XAI may simultaneously increase awareness of, or exposure to, technology-related risks.

Auditor professional-judgment quality had a strong positive effect on operational and economic outcomes, $\beta = 0.781$, and a strong positive association with negative outcomes and risks, $\beta = 0.873$. Its effect on professional credibility and accountability outcomes was not significant. Contrary to the anticipated positive direction, its relationship with audit-quality outcomes was significant and negative, $\beta = -0.159$.

Taken together, the structural results demonstrate that technological maturity, professional and regulatory readiness, organizational infrastructure, explainability, and explanation quality are central drivers of the proposed framework. At the same time, several positive relationships with negative outcomes and risks indicate that successful adoption of explainable artificial intelligence does not eliminate implementation risks and may make such risks more visible or salient.

The reported direct-path results support the proposed intermediary roles of auditor trust, XAI explanation quality, and auditor professional-judgment quality at the conceptual level. However, a definitive statistical conclusion regarding mediation requires the reporting of bootstrapped indirect effects, confidence intervals, and the significance of the corresponding indirect pathways.

4. Discussion and Conclusion

The present study developed and validated an explainable auditing framework for assessing financial statement risk through explainable artificial intelligence using a sequential exploratory mixed-methods design. The qualitative phase identified 102 codes that were consolidated into 42 initial codes, 18 organizing codes, and four overarching dimensions: explainable auditing based on explainable artificial intelligence, antecedents, outcomes, and mediating factors. The quantitative phase generally supported the proposed framework, demonstrating satisfactory construct reliability and convergent validity, substantial explanatory power for several endogenous

variables, and strong predictive relevance for most constructs. The results further indicated that national technology-ecosystem maturity, professional and regulatory factors, organizational and human factors, technical and infrastructural factors, and auditee characteristics functioned as important antecedents of explainable auditing. Explainability and interpretability, professional justifiability, model selection and design, applicability in the audit process, professional trust-building, and organizational requirements subsequently influenced auditor trust, explanation quality, professional judgment, and multiple categories of audit outcomes.

One of the principal qualitative findings was that explainable auditing is a multidimensional concept that cannot be reduced to the technical interpretability of an algorithm. The framework indicated that an explainable audit system must provide traceability of the audit process, comprehensible explanations, alignment with audit risks, support for professional judgment, suitable model design, practical applicability, professional acceptance, and appropriate organizational infrastructure. This finding is consistent with the broader XAI literature, which distinguishes intrinsic and post hoc explainability and emphasizes that explanations must be tailored to the characteristics of users and decision contexts [10-12]. It also aligns with the argument that explainability in auditing should establish a meaningful relationship between model outputs, audit objectives, professional reasoning, and regulatory expectations rather than merely displaying technical feature-importance scores [16-18]. Therefore, the qualitative findings support an integrated conception of explainable auditing in which technical transparency is combined with professional defensibility and organizational usability.

The measurement-model results showed that all constructs had Cronbach's alpha, rho_A, and composite-reliability coefficients above the accepted threshold of 0.70, while all AVE values exceeded 0.50. These findings indicate adequate internal consistency and convergent validity. The strong reliability of the constructs suggests that the indicators identified through the meta-synthesis captured coherent aspects of explainable auditing. This finding is important because explainability is often conceptualized inconsistently across disciplines. The satisfactory reliability of the proposed dimensions indicates that concepts such as explanation quality, auditor trust, professional judgment quality, organizational readiness, and technology-ecosystem maturity can be operationalized as empirically distinguishable dimensions. This multidimensional structure is compatible with maturity-based approaches to practical XAI, which stress that effective explainability depends on analysis, evaluation, organizational readiness, and implementation conditions [15]. It is also consistent with the view that the assessment of intelligent systems requires multidisciplinary criteria extending beyond predictive accuracy [19].

Nevertheless, the discriminant-validity results require careful interpretation. Although all HTMT coefficients remained below 0.90, several exceeded the stricter criterion of 0.85, and the Fornell–Larcker condition was not fully satisfied for auditor trust in XAI models, organizational and infrastructural requirements, explainability and interpretability, XAI explanation quality, and auditor professional-judgment quality. These results suggest close conceptual relationships among the central constructs of the model. This overlap is theoretically understandable because transparent explanations may strengthen trust, facilitate professional judgment, and reflect the adequacy of organizational arrangements simultaneously. Previous research similarly indicates that interpretability, perceived reliability, usefulness, and acceptance are closely interrelated in auditors' responses to AI systems [22, 23]. However, the observed overlap also suggests that the boundaries among these constructs should be refined in subsequent research, particularly between explainability as a system characteristic, explanation quality as an output characteristic, and trust as a user response.

The structural model demonstrated substantial explanatory power. Auditor trust in XAI models had an R^2 value of 0.813, XAI explanation quality had an R^2 of 0.778, negative outcomes and risks had an R^2 of 0.759, auditor

professional-judgment quality had an R^2 of 0.735, and explainability and interpretability had an R^2 of 0.710. These findings indicate that the proposed antecedent and implementation variables explained a large proportion of variance in central dimensions of explainable auditing. Stone–Geisser Q^2 values were also positive and generally above 0.35, confirming strong predictive relevance for most endogenous constructs. This result supports the theoretical proposition that explainable auditing is produced through the interaction of technical, organizational, professional, and contextual conditions rather than through the adoption of a single technological tool. It is consistent with research emphasizing that successful AI implementation in auditing depends on organizational competence, digital capability, regulatory guidance, data access, and the maturity of client information systems [25, 28, 32].

National technology-ecosystem maturity significantly influenced professional trust-building and acceptance, organizational and infrastructural requirements, AI model selection and design, professional justifiability, explainability and interpretability, and applicability in the audit process. Its strongest structural effect was on explainability and interpretability, while the effect-size analysis also showed a large effect of ecosystem maturity on this construct. This finding suggests that explainable auditing depends heavily on the broader technological environment in which audit firms operate. Access to secure infrastructure, suitable software, reliable technology providers, specialist human capital, and high-quality data enables organizations to select more appropriate models and produce more stable and understandable explanations. This result aligns with studies emphasizing that machine-learning-based financial statement risk assessment requires accessible, structured, and traceable data [5]. It is also consistent with evidence that digital capabilities and technological competencies are critical to the sustainability of the auditing profession and its ability to detect fraud risk [25].

Professional and regulatory factors had significant positive relationships with all implementation constructs, including professional trust-building, organizational requirements, model design, professional justifiability, explainability, and applicability. Their particularly strong influence on applicability in the audit process indicates that standards, regulatory recognition, practical guidance, and clarity regarding auditor responsibility are essential for converting explainable AI from an experimental technology into an accepted audit instrument. Auditors may hesitate to use algorithmic explanations if their evidential status is unclear or if professional standards do not specify how such explanations should be validated and documented. This finding is consistent with research emphasizing regulatory transparency, algorithmic accountability, and the need to retain the auditor's ultimate responsibility [27, 30]. It also supports studies showing that explainable AI frameworks are especially important in regulated industries where decisions must be reviewable and defensible [20, 21].

Organizational and human factors significantly affected professional trust-building, organizational requirements, model selection, and applicability in the audit process. Their strongest effect was on professional trust-building and acceptance, indicating that management support, learning culture, interdisciplinary teams, and workforce competencies are central to auditors' willingness to use XAI systems. This finding is aligned with studies showing that auditors require hybrid competencies in accounting, auditing, analytics, and artificial intelligence to interpret model outputs appropriately [23, 25]. It also supports the importance of combining auditor expertise with algorithmic recommendations through human-in-the-loop mechanisms [24]. The negative relationship between organizational and human factors and professional justifiability, however, was unexpected. One possible explanation is that greater emphasis on organizational adoption and human capability may initially expose discrepancies between existing professional reasoning and algorithmic decision processes. As users become more

technically informed, they may evaluate explanations more critically and demand stronger alignment with audit logic.

Technical and infrastructural factors significantly influenced all implementation constructs. This result highlights the importance of valid explanation techniques, integrated data architecture, system stability, metadata quality, and model-validation procedures. The finding is supported by research indicating that explanation methods such as SHAP and LIME can improve the interpretation of complex predictions but must be evaluated for fidelity, consistency, and appropriateness [13]. It also corresponds with studies emphasizing that technically valid explanations require transparent model-processing procedures and suitable criteria for assessing explanation quality [12, 14]. Thus, explainability should not be treated as a descriptive visualization added after model development; it must be incorporated into data governance, system design, testing, and audit documentation.

Auditee characteristics significantly influenced organizational requirements, model selection and design, professional justifiability, explainability and interpretability, and applicability in the audit process. Particularly strong effects were observed on explainability and applicability. This result indicates that the maturity of the client's financial information systems, the effectiveness of internal controls, management cooperation, and data accessibility determine whether explainable auditing can be implemented successfully. Financial statement risk models depend on the quality and completeness of the underlying information, and poor data may produce unreliable predictions and explanations. This finding aligns with studies on financial statement risk assessment and the importance of internal process quality in identifying corporate risk [4]. It also supports the argument that AI-assisted auditing should remain embedded in the auditor's understanding of the entity, its environment, and its internal-control system rather than functioning as an isolated analytical exercise.

Explainability and interpretability significantly enhanced auditor trust, professional credibility and accountability outcomes, operational and economic outcomes, audit-quality outcomes, negative outcomes and risks, XAI explanation quality, and auditor professional-judgment quality. These widespread effects confirm that explainability is a central mechanism in the proposed framework. When auditors can understand the factors underlying a model's output, they can more effectively assess its relevance, challenge its assumptions, document the basis of their conclusions, and integrate its results with other audit evidence. This finding is consistent with research showing that transparent AI can improve the defensibility and accountability of audit judgments [16, 17]. It also supports evidence that AI can strengthen professional judgment when its outputs are integrated with the intelligence, design, and choice stages of decision-making [7, 8].

The positive relationship between explainability and negative outcomes and risks may initially appear contradictory. However, a plausible interpretation is that greater explainability increases the visibility and recognition of technical, cognitive, ethical, and professional risks. Transparent systems may reveal data bias, unstable relationships, problematic assumptions, or ambiguous responsibility that would remain hidden in opaque models. Alternatively, increased exposure to explanations may create an illusion of understanding, leading users to overestimate their comprehension of a complex model. This interpretation is consistent with concerns that simplified explanations may be misleading, that feature importance may be mistaken for causality, and that explanations can create unjustified confidence [26, 27]. Therefore, the positive coefficient should not necessarily be interpreted as evidence that explainability causes harmful outcomes; it may indicate that implementation makes existing risks more observable or introduces new risks requiring governance.

Professional trust-building and acceptance had very strong positive effects on professional credibility and accountability outcomes and operational and economic outcomes, while also improving audit quality, explanation

quality, and professional-judgment quality. These findings suggest that the benefits of XAI are more likely to materialize when auditors perceive the system as credible, fair, ethically aligned, and useful. Professional acceptance can facilitate sustained use, team communication, and organizational learning. The findings align with research showing that perceived usefulness, understanding, competence, and organizational support influence auditors' acceptance of explainable AI [22]. They also support the proposition that AI and robotic process automation can improve audit efficiency when incorporated into structured workflows [1, 2].

Organizational and infrastructural requirements significantly increased auditor trust, operational and economic outcomes, XAI explanation quality, and professional-judgment quality. The strong effect on operational and economic outcomes indicates that clean data, traceable systems, appropriate governance, skilled personnel, and ethical safeguards are necessary to reduce audit time and cost and optimize resource allocation. This finding is compatible with evidence that AI-based learning techniques can improve internal audit cycles and audit processes [33]. It also supports the emphasis placed on data quality, process documentation, and reliable infrastructure in practical AI applications [15]. The positive relationship with negative outcomes and risks again suggests that greater organizational implementation may expose the firm to new responsibilities, privacy concerns, bias risks, or technological dependencies.

AI model selection and design significantly improved auditor trust, professional credibility and accountability outcomes, operational and economic outcomes, explanation quality, and professional-judgment quality. This result indicates that the choice between inherently interpretable models and complex models with post hoc explanations has important professional consequences. Models should be selected according to the audit task, cost of error, data structure, and required level of transparency. The finding aligns with research describing the trade-offs among accuracy, interpretability, and practical utility [10, 11]. It also supports the use of explanation methods and model-validation criteria to ensure that complex predictions remain professionally understandable [13, 14].

XAI explanation quality significantly influenced all four outcome categories. Its strongest effects were on professional credibility and accountability and audit-quality outcomes. This finding confirms that explanations must be clear, relevant, logically consistent, actionable, and aligned with audit findings. Merely producing an explanation is insufficient; the explanation must support the auditor in determining why an item is risky and what procedure should follow. This result is consistent with research emphasizing user-centered explanations and the importance of communication between systems and professional users [18, 32]. It also corresponds with studies arguing that explainable AI can improve regulatory compliance when explanations provide meaningful transparency rather than superficial disclosure [27].

Auditor professional-judgment quality had a strong positive effect on operational and economic outcomes but a significant negative relationship with audit-quality outcomes. The positive operational effect suggests that accurate, flexible, skeptical, and standards-based judgment enables auditors to convert model outputs into efficient audit decisions. However, the negative relationship with audit-quality outcomes may reflect model complexity, suppression effects among highly correlated constructs, or a distinction between perceived judgment quality and objectively measured audit quality. It may also indicate that extensive professional intervention can reduce the consistency or efficiency benefits of algorithmic support when auditors override valid model recommendations. Because the discriminant-validity results showed substantial conceptual proximity among several mediators, this unexpected path should be interpreted cautiously and tested in future studies.

Auditor trust in XAI models significantly improved operational and economic outcomes but did not directly affect professional accountability or audit-quality outcomes. This suggests that trust may primarily facilitate system

use and efficiency, while audit quality and accountability depend more directly on explanation quality, professional acceptance, and interpretability. Trust also positively predicted negative outcomes and risks. This may reflect the danger of automation bias, overreliance, or inappropriate transfer of responsibility to the system. The finding supports human-in-the-loop approaches in which trust remains calibrated and the auditor retains final control [24]. It is also consistent with ethical concerns regarding the weakening of professional responsibility when AI outputs are accepted without critical review [29].

Overall, the findings support a framework in which explainable auditing emerges from the interaction of ecosystem maturity, professional regulation, organizational capability, technical infrastructure, and auditee readiness. These antecedents shape the principal characteristics of explainable auditing, which subsequently affect outcomes through auditor trust, professional judgment, and explanation quality. The results indicate that XAI can improve audit efficiency, risk analysis, accountability, and professional decision-making, but they also demonstrate that explainability does not remove technological, ethical, or cognitive risks. Rather, its value depends on calibrated trust, reliable data, appropriate model design, critical professional judgment, and explicit accountability structures.

The study has several limitations. First, the quantitative phase used a cross-sectional design and a relatively limited sample of 98 auditors, accountants, and information technology specialists; therefore, causal interpretations and broad generalization should be approached cautiously. Second, the data were collected through self-report questionnaires, which may be affected by social-desirability bias, common-method variance, and differences between respondents' perceptions and their actual behavior during audit engagements. Third, several constructs showed high intercorrelations, and the Fornell–Larcker criterion was not fully satisfied for some variables, indicating possible conceptual overlap. Fourth, the model contained numerous simultaneous relationships, which may have contributed to small individual effect sizes and unexpected path directions. Fifth, the qualitative synthesis included publications from different disciplines and methodological traditions, potentially creating variation in the definitions of explainability, trust, judgment quality, and audit outcomes. Finally, indirect effects were not reported; consequently, the proposed mediating roles of auditor trust, explanation quality, and professional-judgment quality could not be conclusively established through formal mediation testing.

Future studies should test the framework using larger and more diverse samples drawn from audit firms, internal audit departments, regulatory agencies, and different national environments. Longitudinal and experimental designs could examine how auditor trust, reliance, and professional judgment change after repeated interaction with explainable AI systems. Researchers should compare different model types and explanation techniques to determine which combinations produce the highest levels of accuracy, comprehension, and professional usefulness. Further studies should refine constructs with high conceptual overlap and evaluate alternative higher-order or hierarchical models. Formal mediation and moderation analyses should be conducted using bootstrapped indirect effects and confidence intervals. Future research should also examine the effects of audit experience, technological competence, firm size, regulatory intensity, task complexity, and client information-system maturity. Behavioral experiments could determine whether explanations reduce or intensify automation bias, confirmation bias, and the illusion of understanding. In addition, objective indicators of audit quality, detection accuracy, audit hours, cost reduction, and review findings should be used alongside perceptual measures.

Audit firms should adopt explainable artificial intelligence through a phased governance program rather than treating it as a standalone analytical technology. The program should establish minimum standards for data quality, model validation, explanation fidelity, bias assessment, cybersecurity, privacy, and documentation. Firms should create interdisciplinary teams composed of auditors, accountants, data scientists, information-system

specialists, and legal or compliance professionals. Auditors should receive applied training in interpreting model outputs, evaluating explanations, recognizing automation bias, and preserving professional skepticism. Each AI-assisted conclusion should retain a documented human review process, and responsibility for the final audit judgment should remain explicitly assigned to the auditor. Firms should select explanation methods according to the audit task and user needs, and explanations should identify influential factors, connect findings to audit risks, and recommend appropriate subsequent procedures. Pilot implementation should begin with well-structured and low-risk applications before expanding to complex judgments. Regulators and professional institutions should also develop practical guidance for the validation, documentation, review, and permissible use of XAI-generated evidence in financial statement audits.

Authors' Contributions

Authors equally contributed to this article.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

Acknowledgments

Authors thank all participants who participate in this study.

Conflict of Interest

The authors report no conflict of interest.

Funding/Financial Support

According to the authors, this article has no financial support.

References

[1] M. Eulerich, J. Pawlowski, N. J. Waddoups, and D. A. Wood, "A Framework for Using Robotic Process Automation for Audit Tasks," *Contemporary Accounting Research,* vol. 39, no. 1, pp. 691-720, 2022, doi: 10.1111/1911-3846.12723.

[2] M. Kassar and M. Jizi, "Artificial Intelligence and Robotic Process Automation in Auditing and Accounting: A Systematic Literature Review," *Journal of Applied Accounting Research,* 2025, doi: 10.1108/JAAR-05-2024-0175.

[3] M. Bozorgasl, P. Pourmahdi, and E. Karimi Qarehomar, "Artificial Intelligence in Auditing Processes," *Certified Accountant,* no. 67, pp. 27-35, 2024.

[4] M. Morshedzadeh, N. Samiei, and F. Mashaiekh, "Assessing Corporate Financial Risk Based on Financial Statement Information," in *First National Conference on Development and Future Studies in Industries with an Internal Processes Approach*, Tehran, 2023. [Online]. Available: https://civilica.com/doc/2035672.

[5] E. Zavitsanos, E. Spyropoulou, G. Giannakopoulos, and G. Paliouras, "Machine Learning for Identifying Risk in Financial Statements: A Survey," *ACM Computing Surveys,* vol. 57, no. 9, p. Article 220, 2025, doi: 10.1145/3723157.

[6] I. Anisykurlillah, H. Mukhibad, K. W. Jati, F. W. Rizkyana, A. Nurkhin, and B. B. Hapsoro, "Audit Committee Attributes and Financial Statement Fraud Risk: Evidence from Indonesian Banks," *Cogent Business & Management,* vol. 12, no. 1, p. 2566443, 2025, doi: 10.1080/23311975.2025.2566443.

[7] A. Sehat Barmacheh, H. Shamakhy, and A. Yalfani, "The Effect of Factors Based on Artificial Intelligence Components (Intelligence, Design, and Choice) on Auditors' Judgment," *Management Accounting and Auditing Knowledge,* vol. 16, no. 61, pp. 265-278, 2025.

[8] A. Sehat Barmacheh, H. Shamakhy, and A. Yalfani, "Presenting an Artificial Intelligence-Based Audit Judgment Model," *Management Accounting and Auditing Knowledge,* vol. 15, no. 60, pp. 191-210, 2026.

[9] V. Ariany, "The Impact of Artificial Intelligence on Audit Quality and Auditor Judgement: A Multi Country Analysis," *Accounting Studies and Tax Journal (COUNT),* vol. 2, no. 3, pp. 557-569, 2025.
[10] K. P. Kalyanathaya and K. K. Prasad, "A Literature Review and Research Agenda on Explainable Artificial Intelligence (XAI)," *International Journal of Applied Engineering and Management Letters (IJAEML),* vol. 6, no. 1, pp. 43-59, 2022, doi: 10.47992/IJAEML.2581.7000.0119.
[11] F. Mumuni and A. Mumuni, "Explainable Artificial Intelligence (XAI): From Inherent Explainability to Large Language Models," *arXiv,* 2025. [Online]. Available: https://arxiv.org/abs/2501.09967.
[12] Q. Lyu and S. M. Wu, "Explainable Artificial Intelligence for Business and Economics: Methods, Applications and Challenges," *Expert Systems,* vol. 42, p. e70017, 2025, doi: 10.1111/exsy.70017.
[13] A. M. Salih *et al.*, "A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME," *Advanced Intelligent Systems,* vol. 7, no. 1, p. 2400304, 2024, doi: 10.1002/aisy.202400304.
[14] B. Liu, P. Liu, W. Lu, and T. Olofsson, "Explainable Artificial Intelligence (XAI) for Material Design and Engineering Applications: A Quantitative Computational Framework," *International Journal of Mechanical System Dynamics,* vol. 5, pp. 236-265, 2025, doi: 10.1002/msd2.70017.
[15] J. Munoz-Ordonez, C. Cobos, J. C. Vidal-Rojas, and F. Herrera, "A Maturity Model for Practical Explainability in Artificial Intelligence-Based Applications: Integrating Analysis and Evaluation (MM4XAI-AE) Models," *International Journal of Intelligent Systems,* 2025, doi: 10.1155/int/4934696.
[16] C. Zhong and S. Goel, "Transparent AI in Auditing Through Explainable AI," *Current Issues in Auditing,* vol. 18, no. 2, pp. A1-A14, 2024, doi: 10.2308/CIIA-2023-009.
[17] V. Golubovskij, "Explainability and Transparency of AI in Auditing," University of Twente, 2024.
[18] E. Edward, "Explainable AI Frameworks for Enhancing Auditor Interpretability and Regulatory Compliance," ed, 2025.
[19] M. Langer, K. Baum, K. Hartmann, S. Hessel, T. Speith, and J. Wahl, "Explainability Auditing for Intelligent Systems: A Rationale for Multi-Disciplinary Perspectives," ed, 2024.
[20] M. Ramos and D. Esther, "Developing Explainable AI Frameworks for Regulated Industries," ed: ResearchGate, 2024.
[21] A. Kesari, D. Sele, E. Ash, and S. Bechtold, "A Legal Framework for Explainable Artificial Intelligence," in "Center for Law & Economics Working Paper Series," ETH Zurich, 2024. [Online]. Available: https://doi.org/10.2139/ssrn.4972085
[22] L. C. Thanh, V. T. M. Duyen, and N. T. T. Hang, "Explainable Artificial Intelligence in Auditing: Factors Influencing Auditors' Acceptance in Vietnam," *Edelweiss Applied Science and Technology,* vol. 9, no. 12, pp. 81-92, 2025, doi: 10.55214/2576-8484.v9i12.11281.
[23] M. Adnan Hamoud, P. Piri, and A. Ashtab, "Feasibility of Using New Artificial Intelligence Technologies to Improve Auditing Processes in the Country," *Accounting and Auditing Review,* vol. 32, no. 3, pp. 535-559, 2025.
[24] A. Taofeek, "Human-in-the-Loop Approaches for Combining AI Recommendations with Auditor Expertise," ed, 2025.
[25] F. M. Razali, N. Sulaiman, D. I. Abdul Manan, and J. Said, "Sustainability of Audit Profession in Digital Technology Era: The Role of Competencies and Digital Technology Capabilities to Detect Fraud Risk," *SAGE Open,* vol. 15, no. 1, p. 21582440241304974, 2025, doi: 10.1177/21582440241304974.
[26] G. Carloni, A. Berti, and S. Colantonio, "The Role of Causality in Explainable Artificial Intelligence," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 2025, doi: 10.1002/widm.70015.
[27] C. Bagwe, "Explainable AI (XAI) in Compliance Audits: Bridging the Gap Between AI and Regulatory Transparency," *International Journal of Scientific Engineering and Science,* vol. 9, no. 3, pp. 137-144, 2025.
[28] E. Bracci, M. Tallaki, and J. E. Otia, "Propensity Factors of Artificial Intelligence Technology Adoption by Public Sector Auditors," *Journal of Public Budgeting, Accounting & Financial Management,* 2025, doi: 10.1108/JPBAFM-09-2024-0189.
[29] N. Noorollahzadeh and M. Sadeghi, "Ethical Consequences of Artificial Intelligence (AI) Adoption in Financial Decision-Making," *Accounting and Management Perspective,* vol. 7, no. 99, pp. 50-60, 2024.
[30] B. Bigpanah, H. Esnaashari, A. Heshi, and G. Asadi, "Accountability of Audit Firms: A Content Analysis Approach," *Accounting and Auditing Review,* vol. 29, no. 2, pp. 213-241, 2022.
[31] M. J. Mansouri, G. Karami, and H. Yazdani, "Identifying Drivers Affecting the Future of Auditing: A Meta-Synthesis Approach," *Accounting and Auditing Review,* vol. 31, no. 2, pp. 390-427, 2024.
[32] J. Dayeh, "Audit and Artificial Intelligence: Audit Data Analytics and Auditing AI," Copenhagen Business School, 2025.
[33] A. Abdi Baraftabi, F. Samadi, S. Shahverdiani, and Z. Hajiha, "The Role of the Internal Audit Cycle Based on Learning Techniques and Artificial Intelligence and Its Effects on Improving Audit Processes," *Management Accounting and Auditing Knowledge,* vol. 16, no. 62, pp. 443-457, 2025.