Auditing at the Algorithmic Frontier: A Critical Narrative Review of Machine Learning and Audit Quality

Zulkiffly Baharom

Tunku Puteri Intan Safinaz School of Accountancy (TISSA-UUM),College of Business, Universiti Utara Malaysia, Malaysia

DOI: https://doi.org/10.51244/IJRSI.2026.1306000006

Received: 21 May 2026; Accepted: 26 May 2026; Published: 17 June 2026

ABSTRACT

The rapid integration of machine learning (ML) and artificial intelligence (AI) into audit practice has generated growing scholarly interest in their implications for audit quality. This narrative review synthesizes 49 peer-reviewed articles sourced from the Web of Science (WoS) database, spanning 2010 to 2026, to critically examine how ML adoption shapes audit quality across diverse institutional and organizational contexts. Drawing on institutional theory and the socio-technical systems framework, this study proposes an integrative conceptual framework that positions five independent variables: institutional pressures, technological capabilities, strategic orientation, ethical frameworks, and AI autonomy level, as antecedents of audit quality, mediated by trust, legitimacy, and decision rights, and moderated by governance mechanisms, organizational culture, and institutional environment. The review reveals that whilst ML meaningfully enhances misstatement detection, risk stratification, and processing efficiency, persistent concerns remain regarding auditor over-reliance, algorithmic opacity, accountability displacement, and professional de-skilling. These findings carry significant implications for standard-setters, audit practitioners, and researchers seeking to govern AI responsibly within the auditing profession.

Keywords: Machine learning, audit quality, artificial intelligence, digital auditing, professional skepticism

INTRODUCTION

The auditing profession is undergoing a profound structural transformation driven by the adoption of ML and AI technologies. What was once a labor-intensive, judgment-dependent process is increasingly supported, and in some instances partially supplanted by algorithms capable of processing vast transactional datasets, identifying anomalous patterns, and generating risk assessments with a speed and scale unattainable by human auditors alone (Jones & Free, 2026; Fedyk et al., 2022). Landmark empirical studies have demonstrated the capacity of ML models to predict material misstatements (Parker et al., 2025), classify financial restatements (Hayes & Boritz, 2021), and assess going concern risk (Gu et al., 2026) with accuracy that rivals or exceeds traditional audit approaches. The proliferation of generative AI tools, including large language models (LLMs), further extends these capabilities into unstructured data analysis and client inquiry responses (Tapis et al., 2026; Lu et al., 2025), portending a fundamental redefinition of what audit quality means in a digitally mediated environment.

Yet the growing body of literature on ML and audit quality remains theoretically fragmented, geographically concentrated, and insufficiently critical. Many extant studies privilege technical performance metrics: detection accuracy, precision-recall rates, over the organizational, institutional, and ethical conditions that determine whether algorithmic capabilities genuinely translate into improved audit outcomes (Alkhatib et al., 2026; Dos Santos & Dos Santos, 2025). Critical questions regarding auditor over-reliance on AI recommendations (Commerford et al., 2022), the erosion of professional skepticism, and the opacity of black-box models have received comparatively scant theoretical attention. Furthermore, the bulk of empirical evidence is concentrated in large, technologically advanced audit markets, notably the United States and China, leaving significant gaps in understanding how ML adoption unfolds in emerging economies and specialized audit contexts such as Islamic finance (Raza et al., 2026; Benhayoun et al., 2025). Against this backdrop, the present narrative review pursues four objectives: (1) to critically synthesize the mechanisms through which ML influences audit quality; (2) to theorize the mediating and moderating conditions that shape this relationship; (3) to identify methodological and thematic gaps in the extant literature; and (4) to propose a conceptual framework that can guide future empirical inquiry. By situating ML adoption within broader institutional and organizational frameworks, this review seeks to move beyond techno-optimistic accounts and offer a more nuanced appraisal of AI's transformative potential in auditing.

LITERATURE REVIEW

Theoretical Foundations

Two theoretical traditions provide the scaffolding for this review. Institutional Theory, particularly its neo-institutional variant (DiMaggio & Powell, 1983), explains why audit firms adopt ML technologies not merely for efficiency gains but in response to coercive, normative, and mimetic isomorphic pressures emanating from regulators, professional bodies, and industry peers. Regulatory mandates from bodies such as the Public Company Accounting Oversight Board (PCAOB) and international standard-setters create coercive pressures to demonstrate technological competence, whilst the actions of Big Four firms generate mimetic pressures that propagate ML adoption across the profession (Law & Shen, 2024; Hope et al., 2025). Complementing this, the Socio-Technical Systems (STS) framework (Trist & Bamforth, 1951) foregrounds the interdependence of technological and human elements in organizational systems. Applied to auditing, STS theory cautions against privileging algorithmic capability over the social and professional dimensions of audit practice: auditor judgment, relational trust with clients, and ethical responsibility, suggesting that audit quality is a jointly produced outcome of human-machine collaboration rather than a purely technical achievement (Alkhatib et al., 2026; MacTavish et al., 2026).

Collectively, these frameworks reorient the research question away from “does ML improve audit quality?” towards the more nuanced inquiry: “under what institutional, organizational, and relational conditions does ML adoption lead to enhanced audit quality?” This reorientation motivates the conceptual framework developed in a subsequent section.

ML Applications in Auditing

The literature documents a diverse and rapidly expanding range of ML applications across the audit lifecycle. In the risk assessment phase, supervised learning algorithms, including logistic regression, random forests, gradient boosting, and support vector machines, have demonstrated strong predictive validity in identifying misstatement-prone accounts and client-specific fraud indicators. Parker et al. (2025) reported that gradient-boosting models trained on financial statement data significantly outperform traditional risk models in predicting material misstatements, whilst Hunt et al. (2021) showed that ML-generated misstatement probability scores systematically influence auditor effort allocation, suggesting that algorithmic risk signals are increasingly embedded in audit planning workflows. Similarly, Jiang and Jones (2018) demonstrated the superior predictive accuracy of ML models for corporate distress prediction compared with conventional discriminant analysis.

In the area of fraud and creative accounting detection, ML approaches have proven particularly valuable for processing complex, high-dimensional datasets that exceed the cognitive capacity of human reviewers. Tieu and Tran (2026) integrated the fraud triangle framework with ML classifiers to detect financial misstatements in an emerging-market context, achieving high classification accuracy whilst noting that model performance varied significantly across industry sectors. Blue et al. (2025) likewise developed an ML-based model for predicting creative accounting in emerging economies, concluding that ensemble methods outperform traditional heuristic approaches. Sanchez-Medina et al. (2019) applied AI to assess whether auditors faithfully reflect organizational reality, rather than client interests, revealing systematic patterns of reporting bias amenable to algorithmic detection.

At the level of audit opinion and going concern assessment, recent studies highlight both the promise and the complexity of human-AI collaboration. Gu et al. (2026) demonstrated that the synergistic integration of ML predictions with auditor judgment yields superior going concern assessments compared to either source alone, a finding that underscores the complementarity, rather than substitutability, of human and algorithmic inputs. Lu et al. (2025) advanced this theme by proposing a dual-model LLM agent framework for audit opinion prediction, leveraging the conversational and reasoning capabilities of large language models to navigate unstructured audit evidence. These developments signal a nascent yet consequential shift from ML as a back-office efficiency tool to ML as a front-end decision-support system, with direct implications for audit quality.

ML Adoption and Audit Quality: Direct Effects

Empirical evidence on the direct relationship between ML adoption and audit quality is substantial but mixed. On the positive side, Fedyk et al. (2022) documented that audit firms deploying AI tools exhibit lower restatement rates and smaller absolute discretionary accruals, which they interpret as evidence of quality enhancement driven by improved risk coverage and analytical precision. Rahman et al. (2024) corroborated these findings in a Chinese context, demonstrating that both audit firm-level and client-level AI adoption are associated with reduced audit fees and shorter audit lag, consistent with efficiency gains that do not compromise detection capability. Tan et al. (2025) extended this analysis to examine client-side AI applications, finding that clients with advanced AI infrastructure receive higher-quality audits due to improved data environments and reduced information asymmetry.

Conversely, a strand of the literature raises serious reservations about whether technological capability automatically translates into improvements in audit quality. Commerford et al. (2022), in a seminal experimental study, demonstrated that auditors exhibit a troubling propensity to defer to AI-generated complex estimates, particularly when those estimates are presented with high confidence, undermining the professional skepticism that is foundational to audit quality. This finding is consistent with broader concerns about automation bias, the tendency to uncritically accept algorithmically generated outputs, which has been documented across numerous professional domains. Fotoh and Lorentzon (2023) adopted an explicitly critical perspective, arguing that audit digitalization may widen rather than narrow the audit expectation gap by creating an impression of infallibility that clients and investors attribute to AI-enhanced audits, thereby escalating expectations beyond what any audit process can reliably deliver.

Contextual and Institutional Influences

The relationship between ML adoption and audit quality is demonstrably conditioned by contextual factors. Organizational readiness, encompassing technological infrastructure, human capital capabilities, and leadership commitment, emerges as a critical enabler of effective ML deployment. Hu et al. (2020) identified organizational readiness and top management support as the most significant facilitators of the adoption of AI-enabled auditing techniques, using a fuzzy-rough set methodology to capture the inherent imprecision in organizational capability assessments. Benhayoun et al. (2025) reported that auditors in emerging economies face acute readiness constraints, with deficits in data infrastructure, digital literacy, and regulatory clarity collectively impeding the quality-enhancing potential of AI adoption. These findings echo Abu Huson et al. (2024), whose bibliometric analysis identified institutional environment as a key boundary condition moderating the IT-audit quality relationship across jurisdictions.

Regulatory and governance frameworks constitute another critical moderating layer. Khan et al. (2025) examined AI adoption, audit quality, and integrated financial reporting in Gulf Cooperation Council (GCC) markets, finding that the quality-enhancing effects of AI adoption are stronger in contexts with robust regulatory oversight and mandatory integrated reporting requirements. This finding aligns with institutional theory's predictions that coercive pressures amplify the quality signals associated with technological adoption. Conversely, in environments characterized by regulatory ambiguity, as documented by Sawaya et al. (2026) in emerging economy audit markets, the adoption of AI for consolidated financial statement audits may generate surface-level efficiency gains without commensurately improving substantive audit quality. Cultural and organizational dimensions are equally salient: Seethamraju and Hecimovic (2023) found that audit firm culture, specifically the degree to which senior partners champion experimentation and tolerate failure, substantially influences both the pace and depth of AI adoption.

Specialized and Emerging Contexts

A meaningful subset of the literature addresses ML and audit quality in specialized or underexplored contexts, complicating and enriching the dominant narrative. In Islamic finance, Raza et al. (2026) conducted a pioneering study examining how AI can reimagine the effectiveness of internal Shariah audits in the Malaysian context, arguing that ML tools capable of processing large volumes of Shariah-compliance transactions hold transformative potential for an audit function historically constrained by limited human expertise and manual documentation. However, the authors cautioned that the values-laden nature of Shariah compliance, which requires interpretive judgment grounded in fiqh al-muamalat rather than purely algorithmic pattern recognition, means that AI adoption must be carefully governed to preserve the integrity and legitimacy of the Shariah audit function. This argument extends the STS framework to a culturally specific domain, highlighting the inadequacy of technologically deterministic accounts that overlook the normative and relational dimensions of professional audit contexts.

The COVID-19 pandemic provided a natural experiment that accelerated AI adoption and exposed its implications for quality amid institutional disruption. Albitar et al. (2020) documented significant declines in audit quality during the early pandemic period, attributing these declines in part to the abrupt shift to remote auditing, which constrained physical evidence-gathering procedures. Grassa et al. (2022) provided a practitioner perspective from the UAE, with auditors reporting that digital audit tools partially compensated for physical access limitations but that the quality of substantive testing deteriorated due to data reliability concerns and reduced client cooperation. The public sector context has received comparatively limited attention: Aslan (2021) argued that the evolving competency requirements of public auditors, increasingly encompassing data analytics and AI literacy, pose a structural challenge for audit institutions that have been slow to develop the human capital needed to leverage ML technologies effectively.

Critical Perspectives and Research Gaps

Notwithstanding the substantial body of evidence reviewed above, several critical gaps and conceptual tensions merit explicit recognition. First, the overwhelming concentration of empirical studies in the United States, China, and Western European markets creates a significant geographic bias in current knowledge. The institutional, cultural, and infrastructural conditions that characterize emerging and developing economies, where audit quality deficits may be most consequential for investor protection and economic development, remain insufficiently understood. Second, methodological homogeneity constrains theoretical advance: the majority of empirical studies rely on archival financial data and quantitative ML models, leaving qualitative dimensions of the human-AI audit relationship, including auditors' phenomenological experience of working with AI, client perceptions of AI-enhanced audits, and the social construction of algorithmic authority, largely unexplored. Third, ethical dimensions of ML in auditing remain underdeveloped. Questions of algorithmic fairness, model explainability, data governance, and the professional responsibility of auditors for ML-generated outputs have received growing attention in adjacent disciplines but have only recently begun to penetrate the auditing literature (MacTavish et al., 2026; Fotoh & Lorentzon, 2023). Fourth, longitudinal studies tracking how the quality implications of ML adoption evolve over time, as models are retrained, regulatory requirements mature, and auditor competencies develop, are virtually absent from the literature.

Proposed Conceptual Framework

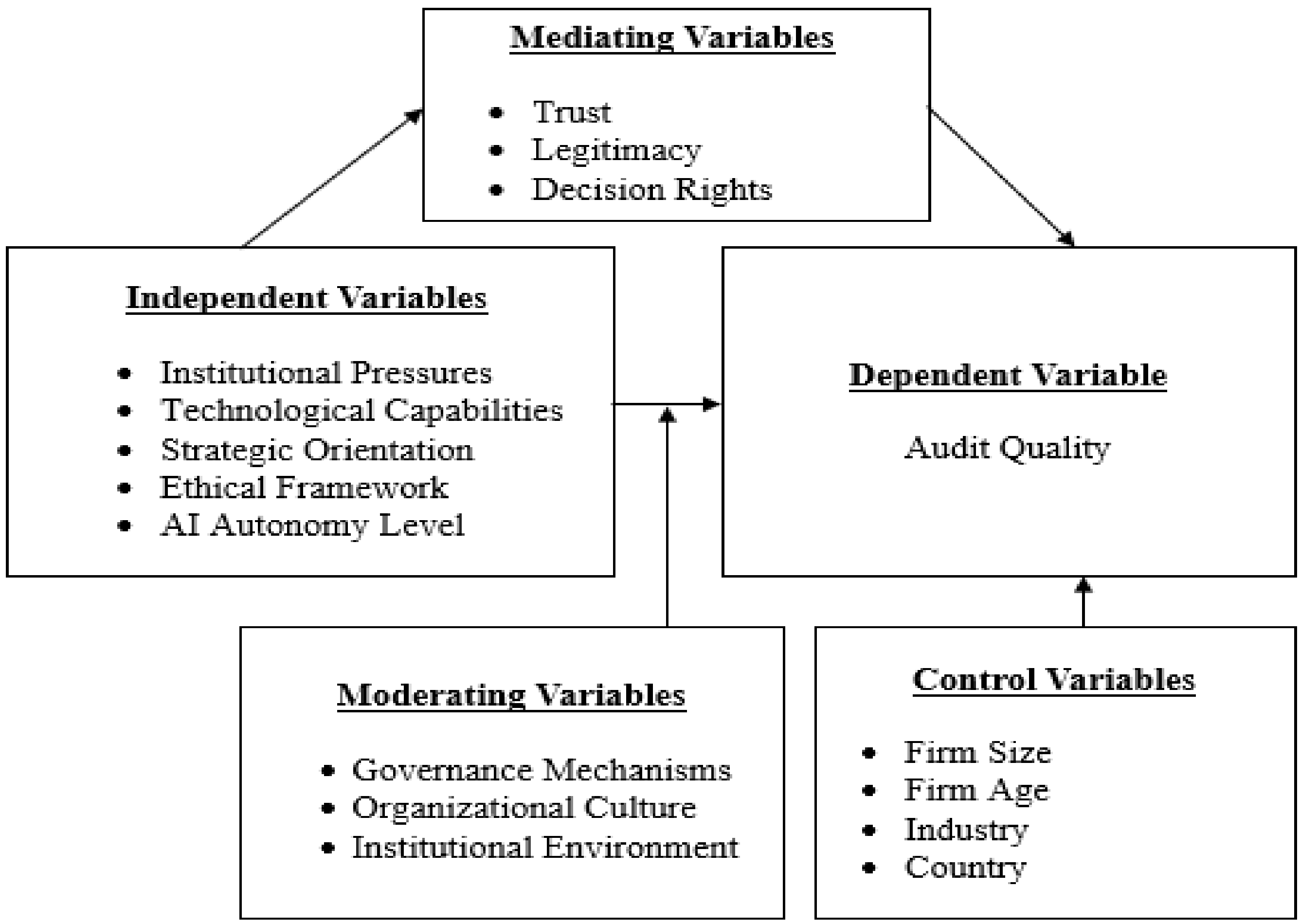
Drawing on the theoretical and empirical review above, this study proposes an integrative conceptual framework that positions ML adoption as a multidimensional organizational phenomenon whose effects on audit quality are neither direct nor uniform but are mediated by relational and institutional mechanisms and moderated by contextual boundary conditions. Table 1 summarises the framework variables, and Figure 1 presents the proposed model.

Table 1: Summary of Conceptual Framework Variables

Role	Variable	Description / Theoretical Basis	Key Sources
Dependent Variable	Audit Quality	The degree to which the audit process conforms to professional standards and detects material misstatements: encompassing accuracy, independence, skepticism, and reliability of audit opinion.	Rahman et al. (2024); Fedyk et al. (2022); Commerford et al. (2022)
Independent Variable 1	Institutional Pressures	Coercive, normative, and mimetic pressures from regulators, professional bodies, and industry peers that drive ML adoption in audit firms.	Law & Shen (2025); Hope et al. (2025); Benhayoun et al. (2025)
Independent Variable 2	Technological Capabilities	The firm's capacity to deploy, maintain, and leverage ML tools encompasses data infrastructure, algorithmic sophistication, and IT investment.	Rahman et al. (2024); Lugli & Bertacchini (2023); Hu et al. (2020)
Independent Variable 3	Strategic Orientation	The degree to which a firm's leadership embraces digital transformation as a strategic priority shapes the depth and quality of AI integration.	MacTavish et al. (2026); Seethamraju & Hecimovic (2023); Manita et al. (2020)
Independent Variable 4	Ethical Framework	The presence of formal governance structures and professional norms that guide the responsible and transparent use of AI in audit engagements.	Jones & Free (2026); Fotoh & Lorentzon (2023); Sanchez-Medina et al. (2019)
Independent Variable 5	AI Autonomy Level	The extent to which ML systems exercise independent analytical judgment in audit tasks, ranging from decision support to near-autonomous processing.	Gu et al. (2026); Tapis et al. (2026); Commerford et al. (2022)
Mediator 1	Trust	The confidence that auditors, clients, and regulators place in AI-generated outputs conditions the degree to which ML recommendations are acted upon.	Chen & Yang (2025); Seethamraju & Hecimovic (2023); Commerford et al. (2022); Noordin et al. (2022)
Mediator 2	Legitimacy	The perceived appropriateness and acceptance of ML use in auditing within professional, regulatory, and societal norms.	Alkhatib et al. (2026); Khan et al. (2025); Fotoh & Lorentzon (2023)
Mediator 3	Decision Rights	The allocation of analytical authority between human auditors and ML systems determines who (or what) holds final accountability for audit judgments.	Gu et al. (2026); MacTavish et al. (2026); Lu et al. (2025)
Moderator 1	Governance Mechanisms	Regulatory frameworks, audit committee oversight, and internal quality control processes that shape the deployment and impact of AI in audit engagements.	Khan et al. (2025); Hope et al. (2025); Hallman et al. (2022)
Moderator 2	Organisational Culture	Shared values, norms, and leadership attitudes within audit firms that determine the receptivity to AI adoption and the cultural governance of its use.	Seethamraju & Hecimovic (2023); Aslan (2021); Manita et al. (2020)
Moderator 3	Institutional Environment	The broader legal, economic, and market conditions. including audit	Sawaya et al. (2026); Raza et al. (2026); Benhayoun et al. (2025)

		market concentration, regulatory maturity, and economic development level, that boundary-condition the ML–quality relationship.	
Control Variable 1	Firm Size	Big N vs. non-Big N firm classification, which affects resource availability for AI investment and the complexity of clients.	Tan et al. (2025); Yang et al. (2020)
Control Variable 2	Firm Age	Years of audit firm operation, proxying institutional inertia, and accumulated professional capital.	Law & Shen (2024)
Control Variable 3	Industry	The client's industry sector, controlling for sector-specific complexity, data availability, and regulatory requirements.	Hunt et al. (2021); Jiang & Jones (2018)
Control Variable 4	Country	National jurisdiction controlling for institutional quality, legal tradition, and technology infrastructure.	Sawaya et al. (2026); Abu Huson et al. (2024)

Figure 1: Proposed Conceptual Framework



Note. The framework integrates institutional theory and socio-technical systems (STS) perspectives. Dashed arrows from moderators indicate boundary-condition effects on the IV→Mediator and Mediator→DV pathways.

To facilitate future empirical testing, indicative operational proxies for each construct are proposed. Audit quality may be operationalized via absolute discretionary accruals, restatement frequency, audit report lag, or going concern accuracy rates. Institutional pressures can be proxied by the number of regulatory enforcement actions, professional body guidance issuances, or Big Four peer adoption rates in the relevant market. Technological capabilities may be measured through firm-level IT capital expenditure, AI-related patent filings, or technology adoption survey indices. Strategic orientation can be captured via the proportion of digital transformation initiatives in firm strategy disclosures or executive survey instruments. Ethical framework strength may be assessed by the presence of formal AI governance policies, code-of-conduct revisions that reference AI, or audit committee mandates for AI oversight. AI autonomy level may be operationalized on a five-point scale ranging from decision-support-only to near-full automation, self-reported by engagement teams or inferred from audit process documentation. These proxies are indicative rather than exhaustive, and researchers are encouraged to adapt them to the methodological demands and data availability of their specific institutional contexts.

METHODOLOGY

Research Design

This study employs a narrative review methodology, a well-established approach in accounting and management scholarship that allows for systematic yet interpretively flexible synthesis of heterogeneous literature (Knopf, 2006). The narrative review approach was selected in preference to systematic or scoping review designs in recognition of the subject matter's inherent conceptual breadth: the relationship between ML and audit quality spans technical, organizational, institutional, and ethical dimensions that resist reduction to a single meta-analytic outcome. Consistent with narrative review standards, the synthesis is guided by explicit inclusion criteria, transparent search documentation, and a commitment to theoretical integration rather than mere descriptive cataloging. The review protocol is reported in alignment with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework (Moher et al., 2009) to ensure methodological transparency.

Search Strategy and Source Selection

A systematic search of the WoS Core Collection was conducted on 21 May 2026 using the following Boolean search string applied to the Topic field: ("auditing quality" OR "audit quality") AND ("artificial intelligence" OR "machine learning"). The initial search yielded 66 records. Sequential filtering was then applied as follows: (i) year filter (2010–2026) retained 66 records; (ii) document type filter (article and review) reduced the corpus to 62 records; (iii) subject area filter (Business Finance, Management, Business, Economics) yielded 50 records; and (iv) language filter (English) produced a final corpus of 49 articles. The year 2010 was chosen as the threshold to capture the period when ML methodologies became computationally tractable for large-scale financial data analysis, following the emergence of scalable ensemble methods and deep learning architectures.

Inclusion and Exclusion Criteria

Articles were included if they: (i) addressed the application of ML or AI to audit processes or audit quality outcomes; (ii) were published in peer-reviewed journals indexed in WoS; (iii) were written in English; and (iv) fell within the defined subject areas and publication window. Articles were excluded if they: (i) addressed AI or ML in non-audit financial contexts without explicit relevance to audit quality; (ii) were conference papers, book chapters, or dissertations; or (iii) lacked sufficient methodological transparency to support critical evaluation. Bibliometric studies that tangentially addressed audit quality were included when they provided substantive theoretical or empirical insights relevant to the review objectives. The final corpus of 49 articles includes studies employing diverse methodologies: archival quantitative analysis, experimental designs, qualitative case studies, bibliometric analyses, and systematic literature reviews, reflecting the field's pluralistic epistemological landscape.

Analytical Approach

Articles were analyzed through a theoretically informed thematic synthesis approach (Thomas & Harden, 2008). Each article was reviewed for: (i) the ML or AI technologies examined; (ii) the audit quality dimensions addressed; (iii) theoretical frameworks invoked; (iv) methodological design; (v) key empirical findings; and (vi) acknowledged limitations. Thematic categories were developed iteratively through constant comparison, with initial categories grounded in prior literature and subsequently refined to accommodate emergent themes from the corpus. The conceptual framework presented in the Literature Review section was derived through iterative synthesis, integrating both the dominant findings and the critical challenges identified throughout the corpus.

FINDINGS

Theme 1: ML as a Quality-Enhancing Technology — Evidence and Limits

The strongest and most consistent finding across the corpus is that ML adoption is positively associated with improvements in technically measurable proxies of audit quality. Studies employing archival financial data consistently document reductions in absolute discretionary accruals, restatement rates, and audit fees, the latter interpreted as a quality-consistent efficiency signal, among clients of AI-adopting audit firms (Tan et al., 2025; Rahman et al., 2024; Fedyk et al., 2022). The superiority of ML classifiers over conventional statistical models in predicting material misstatements and financial distress is well established across multiple algorithms and datasets (Parker et al., 2025; Abdi et al., 2025; Jiang & Jones, 2018), with ensemble methods, particularly gradient boosting and random forests, consistently outperforming single-algorithm approaches.

These academic findings are corroborated by practitioner-level deployments. EY’s Helix platform, KPMG’s Clara, and Deloitte’s Omnia AI suite represent flagship examples of ML integration into audit workflows at scale, enabling near-complete population testing of journal entries, automated anomaly flagging, and real-time risk scoring across large client datasets. Notably, the KPMG Clara Contracts tool has been deployed to analyze lease contracts for IFRS 16 compliance, whilst EY’s Canvas platform applies natural language processing to extract structured data from unstructured client documents, including board minutes and legal agreements. These implementations illustrate that ML in practice operates not as a wholesale replacement for auditor judgment but as a targeted enhancement of evidence-gathering efficiency in well-defined procedural domains, consistent with the STS framework’s emphasis on structured human-machine complementarity.

However, these performance findings require contextualization. The majority of high-performing models are trained and tested on data from well-developed capital markets characterized by robust disclosure environments, standardized financial reporting, and large historical training datasets. As Sawaya et al. (2026) and Benhayoun et al. (2025) document, model performance degrades significantly in emerging-market contexts where data quality is lower, accounting standards are applied less consistently, and audit-trail documentation is incomplete. Furthermore, the construct validity of archival audit quality proxies is itself contested: discretionary accruals, whilst widely used, capture only a narrow slice of audit quality and are subject to well-known measurement error problems. The reliance on these proxies may systematically overstate the quality-enhancing effects of ML adoption whilst obscuring quality dimensions, such as the rigor of substantive testing, the depth of professional skepticism exercised, or the adequacy of client communication, that resist quantification.

Theme 2: Human-AI Collaboration and the Problem of Over-Reliance

One of the most consequential and theoretically rich findings in the corpus concerns the relationship between auditor judgment and AI-generated recommendations. The experimental evidence is unambiguous: auditors exhibit systematic tendencies to anchor on AI outputs, defer to algorithmic estimates, and reduce their independent scrutiny when ML systems signal confidence (Commerford et al., 2022; Hunt et al., 2022). This automation bias, a well-documented cognitive phenomenon in human-computer interaction research, is particularly pronounced when AI recommendations align with auditors' initial assessments, reinforcing rather than challenging professional pre-judgments. Gu et al. (2026) demonstrate that optimal audit outcomes are achievable when human judgment and ML predictions are systematically integrated in a structured collaborative framework, but that this synergy requires deliberate design and does not emerge spontaneously from AI deployment.

The implications for professional skepticism, the cornerstone of audit quality under ISA 200, are profound. If auditors systematically reduce their critical scrutiny in the presence of AI-generated assurances, the quality-enhancing potential of ML technology may be partially or wholly offset by the degradation of the human judgment capacity that underpins independent audit opinion formation. MacTavish et al. (2026) extend this concern to staff-level auditors, demonstrating that junior auditors face acute professional development risks when AI automates the routine analytical tasks that have historically served as the primary vehicle for developing professional judgment. This finding resonates with broader de-skilling concerns in the sociology of work literature and suggests that the long-term workforce implications of ML adoption in auditing may be more problematic than short-term quality metrics indicate.

Theme 3: Institutional and Governance Conditions as Quality Moderators

A robust finding across the corpus is that the quality-enhancing effects of ML adoption are neither universal nor unconditional but are substantially moderated by institutional and governance contexts. Regulatory quality emerges as a particularly potent moderator: in jurisdictions with strong enforcement, mandatory disclosure requirements, and active oversight bodies, ML adoption is associated with more pronounced audit quality improvements, consistent with the institutional theory proposition that coercive pressures amplify quality signals (Khan et al., 2025; Hope et al., 2025). In contrast, in regulatory environments characterized by inconsistent enforcement or weak investor protection, ML adoption may generate performative quality improvements, visible in output metrics, without substantive changes to audit rigor.

Audit market structure constitutes a related moderating force. Yang et al. (2020), using a double machine-learning approach with gradient boosting, provided refined causal evidence for the Big N audit quality premium, demonstrating that the quality gap between Big Four and non-Big Four firms is partially mediated by differential access to AI capabilities. Law and Shen (2024) examined how AI reshapes audit firm structures, finding that AI adoption exacerbates rather than ameliorates quality stratification within the profession, as large firms with greater capacity for AI investment widen their quality advantage over smaller competitors. This finding challenges optimistic narratives that portray AI as a democratising technology capable of elevating quality standards across the profession irrespective of firm size.

Theme 4: Emerging and Specialized Contexts — Divergence from Mainstream Findings

Studies conducted in emerging economies and specialized audit contexts provide important correctives to the mainstream literature's predominantly optimistic assessment. Raza et al. (2026) demonstrate that applying AI to Islamic financial institution auditing presents distinctive challenges rooted in the normative and interpretive demands of Shariah compliance assessment, in which algorithmic classification of complex jurisprudential questions may yield technically accurate yet substantively inadequate audit conclusions. This finding challenges the implicit assumption in much of the ML-audit literature that audit quality is a culturally neutral and universally quantifiable construct, pointing instead to the need for context-sensitive quality frameworks that acknowledge the institutional and normative embeddedness of audit practice.

In COVID-19-affected contexts, Albitar et al. (2020) and Grassa et al. (2022) documented that audit quality deteriorated even when digital tools were available, suggesting that the availability of technology is a necessary but not sufficient condition for maintaining quality amid institutional disruption. The pandemic experience thus provides a natural experiment that reveals the limits of technological substitutes for in-person evidence gathering and the irreplaceable role of auditors' physical access in certain quality-critical procedures. Mugwira (2022) and Abu Huson et al. (2024), in their bibliometric analyses, identify the systematic underrepresentation of African and Middle Eastern research contexts in the AI-audit quality literature and call for deliberate efforts to diversify the geographic base of future research. A parallel blind spot concerns small and mid-tier audit firms, whose resource constraints, limited IT infrastructure, and narrower client bases create a qualitatively different adoption context that the existing literature, dominated by Big Four and large-firm samples, fails to adequately capture. The quality implications of ML adoption for non-Big N firms operating in resource-scarce environments remain an important, largely unexplored frontier.

Theme 5: Ethical, Accountability, and Expectation Gap Dimensions

A growing but still nascent strand of the literature addresses the ethical and accountability dimensions of ML in auditing. Fotoh and Lorentzon (2023) advanced a critical argument that audit digitalization creates a new dimension of the audit expectation gap: users of financial statements may attribute an unjustified level of infallibility to AI-enhanced audits, generating expectations that exceed the inherent limitations of any probabilistic system. This 'algorithmic expectation gap', distinct from the traditional performance expectation gap, represents a novel quality risk that standard-setters and audit firms have yet to adequately address. Sanchez-Medina et al. (2019) raised related concerns about the potential for AI systems to embed or amplify existing biases in audit judgments, particularly where training data reflects historically biased reporting patterns.

The question of accountability for ML-generated audit errors, who bears professional and legal responsibility when an algorithm-assisted audit fails to detect a material misstatement, remains largely unresolved in both the legal and professional standards literature. Jones and Free (2026) argue that accounting practitioners are insufficiently prepared to navigate these accountability questions, calling for substantial reforms to professional education and continuing development programs. Alkhatib et al. (2026), drawing on STS theory, emphasize that the key accountability risk is the displacement of human responsibility onto algorithmic systems, a displacement that may be actively cultivated by audit firms seeking to insulate themselves from liability whilst simultaneously exploiting the marketing value of AI-enhanced audit claims.

DISCUSSION

Theoretical Implications

The findings of this review carry significant implications for the theoretical understanding of AI in professional service organizations. From an institutional theory perspective, the evidence confirms that ML adoption in auditing is driven primarily by isomorphic pressures rather than purely rational efficiency calculations, with regulatory mandates, professional body guidance, and peer-firm behavior exerting considerable influence on adoption trajectories. This insight has important consequences for predictions about the quality effects of adoption: firms that adopt ML primarily to signal legitimacy rather than to genuinely enhance analytical capability may exhibit performance improvements on observable quality metrics whilst failing to internalize the deeper organisational changes, in workforce capability, quality culture, and decision-making architecture, necessary for substantive quality improvement. The proposed framework's incorporation of legitimacy as a mediating variable is designed to capture precisely this distinction.

The STS framework's emphasis on the co-constitution of human and technical agency provides equally valuable theoretical traction. The consistent finding that optimal audit quality emerges from structured human-AI collaboration, rather than from algorithmic autonomy, is fully consonant with the STS proposition that technological capability is most effectively realized when it is embedded in appropriately configured social systems. Audit firms that deploy ML tools without investing in the human capital, governance structures, and professional culture necessary to manage human-AI interaction constructively are likely to realize sub-optimal quality returns. This theoretical insight also suggests that research designs that treat ML adoption as a binary variable, adopted versus non-adopted, are likely to underestimate the variance in quality outcomes attributable to differences in adoption quality and organizational embeddedness.

Practical Implications

For audit practitioners and firm leadership, the findings underscore the imperative to invest in auditor education and professional development that specifically address the cognitive and ethical challenges of working with AI systems. Training programs that cultivate critical evaluation of algorithmic outputs, rather than deferential acceptance, are essential to preserve the professional skepticism that remains the bedrock of audit quality. The evidence on automation bias (Commerford et al., 2022) suggests that this training must be active and experiential, exposing auditors to scenarios in which AI recommendations are misleading, rather than relying on declarative instruction about the limits of algorithmic systems.

For standard-setters and regulators, the findings call for urgent attention to several governance lacunae. The absence of explicit regulatory guidance on auditor responsibility for ML-generated outputs creates an accountability vacuum that neither serves investor protection nor provides practitioners with clear professional guidance. The PCAOB, IAASB, and national audit oversight bodies should consider developing technology-specific quality standards that address model validation requirements, documentation obligations for algorithmic audit decisions, and auditor responsibility for AI-assisted judgments. Regulators in emerging economies face the additional challenge of establishing enabling infrastructure, data quality standards, digital audit trail requirements, and AI literacy benchmarks for audit professionals before quality-enhancing ML adoption can be meaningfully pursued.

Particular attention should be directed to the governance of explainability. The opacity of many high-performing ML models, particularly deep learning architectures and ensemble methods, creates a fundamental tension with the audit profession's obligation to document and justify the basis for audit conclusions. Regulators and standard-setters should consider mandating the use of eXplainable AI (XAI) techniques, such as SHAP (SHapley Additive exPlanations) value analysis or LIME (Local Interpretable Model-agnostic Explanations), in audit contexts where ML outputs influence material audit judgments, so that algorithmic reasoning can be inspected, challenged, and documented in audit working papers. More broadly, an ethical audit governance framework, encompassing principles of algorithmic transparency, human oversight thresholds, bias monitoring, and stakeholder accountability, is urgently needed to provide professional and regulatory coherence to a landscape currently characterized by ad hoc firm-level policies and the absence of enforceable standards.

Limitations

This review is subject to several limitations that qualify the generalisability of its conclusions. As a narrative review, it is inherently susceptible to selection and interpretation bias, despite the structured search protocol and thematic synthesis approach employed. The restriction to English-language WoS-indexed articles may have excluded relevant contributions published in other languages or databases, particularly from non-English-speaking emerging markets. The rapidly evolving nature of ML technology means that some findings, particularly those related to specific algorithms or tool capabilities, may have limited shelf life as the technological frontier advances. Finally, the proposed conceptual framework, whilst grounded in the reviewed literature, remains empirically untested; its predictive validity is contingent on future quantitative and qualitative research explicitly designed to examine the mediated moderation pathways it theorizes.

CONCLUSION

Summary of Contributions

This narrative review has synthesized 49 peer-reviewed studies to develop a critical, theoretically grounded understanding of the relationship between machine learning adoption and audit quality. The review's central contribution is the demonstration that this relationship is neither straightforward nor universally positive, but is contingent on mediating mechanisms, trust, legitimacy, and the allocation of decision rights between human auditors and AI systems, and is bounded by moderating conditions, including governance quality, organizational culture, and the institutional environment. By integrating institutional theory and the socio-technical systems framework, this review provides a theoretically coherent account of why ML adoption produces divergent quality outcomes across different firms, jurisdictions, and audit contexts.

Future Research Directions

The review identifies several high-priority directions for future research. First, longitudinal studies tracking the quality implications of ML adoption over extended time horizons are urgently needed to assess whether initial quality improvements are sustained as models age, regulatory environments mature, and auditor competencies evolve. Second, qualitative and mixed-methods research designs are required to illuminate the phenomenological and social dimensions of human-AI interaction in audit practice, including auditors' lived experiences, professional identity responses to AI adoption, and client perceptions of algorithmic audit quality. Third, comparative cross-national studies that explicitly theorize institutional environment as a moderating variable, rather than treating it as background noise, would substantially advance understanding of how contextual factors shape the ML-quality relationship. Fourth, research on Islamic finance, public sector, and SME audit contexts is warranted to extend the geographic and institutional scope of current knowledge beyond the predominantly large-firm, advanced-economy focus of the existing literature. Fifth, ethical and governance-focused research is needed to develop frameworks for auditor accountability in AI-assisted audit environments, addressing the algorithmic expectation gap and the de-skilling risks documented in this review. Collectively, these research directions chart a path towards a more comprehensive, critical, and practically relevant understanding of artificial intelligence as a transformative, but not unconditionally beneficial, force in the auditing profession.

REFERENCES

1. Abdi, M., Moslemi, A., & Rashidi, M. (2025). A machine learning approach to assessing audit quality (AQ) in company with non-switching auditors: Extra trees classifier (ETC) model. Interdisciplinary Journal of Management Studies, 19(1), 121–135. https://doi.org/10.22059/ijms.2025.384690.677133
2. Abu Huson, Y., Sierra-García, L., & Garcia-Benau, M. A. (2024). A bibliometric review of information technology, artificial intelligence, and blockchain on auditing. Total Quality Management & Business Excellence, 35(1–2), 91–113. https://doi.org/10.1080/14783363.2023.2256260
3. Albitar, K., Gerged, A. M., Kikhia, H., & Hussainey, K. (2020). Auditing in times of social distancing: the effect of COVID-19 on auditing quality. International Journal of Accounting and Information Management, 29(1), 169–178. https://doi.org/10.1108/ijaim-08-2020-0128
4. Alkhatib, E., Alkhatib, A., & Jarvis, R. (2026). Auditing using artificial intelligence: A systematic literature review with scientometric and topic modeling insights through a socio-technical systems (STS) perspective. Journal of Financial Reporting and Accounting, 1–28. https://doi.org/10.1108/jfra-07-2025-0518
5. Aslan, L. (2021). The evolving competencies of the public auditor and the future of public sector auditing. In Contemporary Issues in Public Sector Accounting and Auditing (pp. 113–129). Emerald Publishing Limited. https://doi.org/10.1108/S1569-375920200000105008
6. Benhayoun, I., Bougrine, S., & Sassioui, A. (2025). Readiness for artificial intelligence adoption by auditors in emerging countries – a PLS-SEM analysis of Moroccan firms. Journal of Financial Reporting and Accounting, 23(4), 1486–1508. https://doi.org/10.1108/jfra-07-2024-0448
7. Blue, G., Chahrdahcheriki, M., Rezaee, Z., & Khotanlou, M. (2025). A model for predicting creative accounting in emerging economies. International Journal of Accounting and Information Management, 33(1), 1–31. https://doi.org/10.1108/ijaim-09-2023-0240
8. Chen, S., & Yang, J. (2025). Intelligent manufacturing, auditor selection and audit quality. Management Decision, 63(3), 964–997. https://doi.org/10.1108/md-09-2023-1518
9. Commerford, B. P., Dennis, S. A., Joe, J. R., & Ulla, J. W. (2022). Man versus machine: Complex estimates and auditor reliance on artificial intelligence. Journal of Accounting Research, 60(1), 171–201. https://doi.org/10.1111/1475-679x.12407
10. DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147. https://doi.org/10.2307/2095101
11. Dos Santos, D. L. D. S., & Dos Santos, G. C. (2025). Technological convergence in financial auditing: A systematic literature review. Data Science in Finance and Economics, 5(4), 440–465. https://doi.org/10.3934/dsfe.2025018
12. Fedyk, A., Hodson, J., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? Review of Accounting Studies, 27(3), 938–985. https://doi.org/10.1007/s11142-022-09697-x
13. Fotoh, L. E., & Lorentzon, J. I. (2023). Audit digitalization and its consequences on the audit expectation gap: A critical perspective. Accounting Horizons, 37(1), 43–69. https://doi.org/10.2308/horizons-2021-027
14. Grassa, R., Obaidalla, I., & Hamza, M. (2022). Auditors' perspective of audit quality during the COVID-19 pandemic: Evidence from the United Arab Emirates. Indonesian Journal of Sustainability Accounting and Management, 6(2), 302–313. https://doi.org/10.28992/ijsam.v6i2.623
15. Gu, Y., Parker, C. A. Z., & Vasarhelyi, M. (2026). Synergizing machine and human judgment: evidence from going concern assessment. Journal of Accounting and Public Policy, 57(107435). https://doi.org/10.1016/j.jaccpubpol.2026.107435
16. Hallman, N. J., Kartapanis, A., & Schmidt, J. J. (2022). How do auditors respond to competition? Evidence from the bidding process. Journal of Accounting and Economics, 73(2–3), 101475. https://doi.org/10.1016/j.jacceco.2021.101475

17. Hayes, L., & Boritz, J. E. (2021). Classifying restatements: An application of machine learning and textual analytics. Journal of Information Systems, 35(3), 107–131. https://doi.org/10.2308/isys-19-003
18. Hope, O. K., Wang, C., Wu, Y., & Zhang, M. (2025). Does convergence with International Standards on auditing improve audit quality? The Accounting Review, 100(2), 189–218. https://doi.org/10.2308/tar-2022-0610
19. Hu, K. H., Chen, F. H., Hsu, M. F., & Tzeng, G. H. (2020). Identifying key factors for adopting artificial intelligence-enabled auditing techniques by joint utilization of fuzzy-rough set theory and mrdm technique. Technological and Economic Development of Economy, 27(2), 459–492. https://doi.org/10.3846/tede.2020.13181
20. Hunt, E., Hunt, J., Richardson, V. J., & Rosser, D. (2022). Auditor response to estimated misstatement risk: A machine learning approach. Accounting Horizons, 36(1), 111–130. https://doi.org/10.2308/horizons-19-139
21. Hunt, J. O. S., Rosser, D. M., & Rowe, S. P. (2021). Using machine learning to predict auditor switches: How the likelihood of switching affects audit quality among non-switching clients. Journal of Accounting and Public Policy, 40(5), 106785. https://doi.org/10.1016/j.jaccpubpol.2020.106785
22. Jiang, Y., & Jones, S. (2018). Corporate distress prediction in China: a machine learning approach. Accounting and Finance, 58(4), 1063–1109. https://doi.org/10.1111/acfi.12432
23. Jones, S., & Free, C. (2026). What accountants need to know about artificial intelligence and machine learning: a review and call for future research. Journal of Accounting Literature, 48(5), 133–158. https://doi.org/10.1108/jal-12-2025-0693
24. Khan, F., Ullah Jan, S., & Zia-ul-haq, H. M. (2025). Artificial intelligence adoption, audit quality and integrated financial reporting in GCC markets. Asian Review of Accounting, 33(3), 464–495. https://doi.org/10.1108/ara-03-2024-0085
25. Knopf, J. W. (2006). Doing a literature review. PS: Political Science & Politics, 39(1), 127–132.
26. Law, K. K. F., & Shen, M. (2024). How does artificial intelligence shape audit firms? Management Science, 71(5), 3641–3666. https://doi.org/10.1287/mnsc.2022.04040
27. Lu, Y., Hao, J., & Tang, X. (2025). Dual-model synergy for audit opinion prediction: A collaborative LLM agent framework approach. International Review of Economics & Finance, 104(104642). https://doi.org/10.1016/j.iref.2025.104642
28. Lugli, E., & Bertacchini, F. (2023). Audit quality and digitalization: some insights from the Italian context. Meditari Accountancy Research, 31(4), 841–860. https://doi.org/10.1108/medar-08-2021-1399
29. MacTavish, C., Fiolleau, K., Osecki, E., & Thorne, L. (2026). Technology and its implications for Staff Auditors. Accounting Horizons, 40(1), 103–115. https://doi.org/10.2308/horizons-2023-057
30. Manita, R., Elommal, N., Baudier, P., & Hikkerova, L. (2020). The digital transformation of external audit and its impact on corporate governance. Technological Forecasting and Social Change, 150(119751). https://doi.org/10.1016/j.techfore.2019.119751
31. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Medicine, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097
32. Mugwira, T. (2022). Internet-related technologies in the auditing profession: A WOS bibliometric review of the past three decades and conceptual structure mapping. Revista de Contabilidad, 25(2), 201–216. https://doi.org/10.6018/rcsar.428041
33. Noordin, N. A., Hussainey, K., & Hayek, A. F. (2022). The use of artificial intelligence and audit quality: An analysis from the perspectives of external auditors in the UAE. Journal of Risk and Financial Management, 15(8), 339. https://doi.org/10.3390/jrfm15080339
34. Parker, C. Z., Jiang, L., Cho, S., & Vasarhelyi, M. A. (2025). Predicting material misstatements using machine learning. The Accounting Review, 100(6), 225–262. https://doi.org/10.2308/tar-2024-0035
35. Rahman, M. J., Zhu, H., & Yue, L. (2024). Does the adoption of artificial intelligence by audit firms and their clients affect audit quality and efficiency? Evidence from China. Managerial Auditing Journal, 39(6), 668–699. https://doi.org/10.1108/maj-03-2023-3846
36. Raza, M., Sori, Z. M., & Shamsher, M. (2026). Reimagining the effectiveness of internal Shariah audit through artificial intelligence: the Malaysian context. Journal of Islamic Accounting and Business Research, 1–16. https://doi.org/10.1108/jiabr-05-2025-0307

37. Sánchez-Medina, A. J., Blázquez-Santana, F., & Alonso, J. B. (2019). Do auditors reflect the true image of the company contrary to the clients' interests? An artificial intelligence approach. Journal of Business Ethics, 155(2), 529–545. https://doi.org/10.1007/s10551-017-3496-4
38. Sawaya, C., Khalil, F. C., Jabbour Al Maalouf, N., & Kaspard, J. (2026). Revolutionizing audit practice in emerging economies: the role of Artificial Intelligence in enhancing the audit of consolidated financial statements. Cogent Business & Management, 13(1). https://doi.org/10.1080/23311975.2026.2625974
39. Seethamraju, R., & Hecimovic, A. (2023). Adoption of artificial intelligence in auditing: An exploratory study. Australian Journal of Management, 48(4), 780–800. https://doi.org/10.1177/03128962221108440
40. Tan, J., Chang, S., Zheng, Y., & Chan, K. C. (2025). Effects of artificial intelligence in the modern business: Client artificial intelligence application and audit quality. International Review of Financial Analysis, 104(104271). https://doi.org/10.1016/j.irfa.2025.104271
41. Tapis, G. P., Ravenscraft, J., Naegle, J. C., Jr, Keller, C. E., Jr, & Church, K. S. (2026). ChatGPT and the financial statement audit: Can staff level employees at audit clients use ChatGPT to answer auditor inquiries? Journal of Emerging Technologies in Accounting, 1–13. https://doi.org/10.2308/jeta-2024-040
42. Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8(1), 45. https://doi.org/10.1186/1471-2288-8-45
43. Tieu, T. T. H., & Tran, N. H. (2026). Integrating the fraud triangle with machine learning for financial misstatement detection: Evidence from an emerging market. Cogent Business & Management, 13(1). https://doi.org/10.1080/23311975.2026.2614367
44. Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting: An examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system. Human Relations; Studies Towards the Integration of the Social Sciences, 4(1), 3–38. https://doi.org/10.1177/001872675100400101
45. Yang, J. C., Chuang, H. C., & Kuan, C. M. (2020). Double machine learning with gradient boosting and its application to the Big N audit quality effect. Journal of Econometrics, 216(1), 268–283. https://doi.org/10.1016/j.jeconom.2020.01.018