

Systematic Review

Artificial Intelligence and Machine Learning in the Transformation of Forensic Accounting: A Systematic Literature Review

Luis Ángel Meneses Cerón ^{1,*}, Eliana Fernández Rengifo ¹, Victoria Eugenia Pino Terán ^{2,*},
María Alejandra Albán Lara ², Rosa Cristina Vega Sánchez ¹ and Jhon Jairo Fuentes Sánchez ³

¹ Faculty of Accounting, Economic, and Administrative Sciences, Universidad del Cauca, Popayán 190001, Colombia; elianafernandez@unicauca.edu.co (E.F.R.); rosavega@unicauca.edu.co (R.C.V.S.)

² Faculty of Business Sciences, Corporación Universitaria ComfacaUCA-UnicomfacaUCA, Popayán 190001, Colombia; malban@unicomfacaUCA.edu.co

³ Faculty of Administrative and Accounting Science, Universidad Cooperativa de Colombia, Barrancabermeja 687031, Colombia; jhonj.fuentes@ucc.edu.co

* Correspondence: angelm@unicauca.edu.co (L.Á.M.C.); vpino@unicomfacaUCA.edu.co (V.E.P.T.)

Abstract

This research analyzed the integration of artificial intelligence (AI) and machine learning in forensic accounting as tools for the detection of accounting fraud. Through a systematic review of the literature in Scopus under the PRISMA protocol, 76 documents published between 2019 and 2026 were selected and analyzed. The results show that supervised machine learning techniques, assembly methods, deep learning and natural language processing have demonstrated superior capacity to rule-based approaches to detect fraudulent patterns, when quality data and professionals capable of interpreting their results with forensic criteria are available. A recurring finding throughout the corpus is that machine learning models tend to perform better than rule-based approaches in contexts where sufficient quality data is available. Although this advantage is not universal, it depends on the type of fraud analyzed, the design of the model, the evaluation metrics used, and the institutional context of implementation. AI expands the practitioner's ability to identify anomalous patterns in volumes of information that manual analysis cannot cover. However, significant barriers persist related to data quality, model interpretability, professional resistance to technological updating, and the absence of regulatory frameworks that grant evidentiary validity to algorithmic results. Latin America and Africa show a low investigative production in this field, despite concentrating high levels of vulnerability to financial fraud. The study concludes that the transformation of forensic accounting using AI is a social, institutional and educational process, not exclusively technical, which requires interdisciplinary training, data governance and specific ethical frameworks for the forensic field.



Academic Editors: Naser Makarem
and Hassan Yazdifar

Received: 30 April 2026

Revised: 17 June 2026

Accepted: 23 June 2026

Published: 10 July 2026

Copyright: © 2026 by the authors.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: artificial intelligence; machine learning; forensic accounting; fraud detection; systematic literature review

1. Introduction

Financial fraud is not a new phenomenon, but it is one that has found increasingly fertile ground in the digital age. The accelerated digitalization of business, the globalization of markets, and the growing complexity of transactions have exposed an uncomfortable reality: many companies, institutions, and economic agents operate with control gaps that make them vulnerable to irregularities of various kinds (Erazo-Castillo & De la A-Muñoz,

2023). This vulnerability is not always due to carelessness; to a large extent, it reflects the fact that traditional supervisory mechanisms are no longer sufficient to cover the speed and volume of current operations (Sun et al., 2026).

Faced with this scenario, forensic accounting has gained relevance as a specialized field that articulates accounting, auditing, and investigation knowledge to detect fraud, identify discrepancies, and build valid technical evidence in judicial processes (Kanchan, 2021). Its scope goes beyond the accounting record, for Renzhou (2011), it incorporates financial valuation and technological tools to address problems of illegal appropriation, patrimonial damage and changes in the value of assets, with the purpose of safeguarding the assets of entities through objective and traceable expert reports that can be sustained both in civil and criminal proceedings.

Financial fraud, in this context, is understood as a set of deceptive practices through which a person or an organization manipulates or distorts economic information to obtain an undue benefit, exploiting the trust of those who manage resources or report data (Fu et al., 2025). Its seriousness in the accounting field lies precisely in the fact that it erodes the transparency of financial statements, affecting the decisions of investors, creditors, and other users who depend on that information to act (IASB, 2018).

It is at this point that AI and emerging technologies have begun to substantially transform accounting and financial supervision processes, from manual procedures to systems based on advanced software, which have reduced the time spent on repetitive tasks, reduced human error, and improved the ability to respond to risk situations (Tavares et al., 2023; Lehner & Knoll, 2022). Applying these anti-corruption methodologies well is key to reducing financial irregularities because it allows accounting capabilities to be complemented with modern techniques that uncover fraud schemes and prevent deceptive practices through the automated analysis of large volumes of data (Cappa et al., 2024). However, although the applications of AI in accounting have decades of history, ranging from auditing to financial reporting, its effective integration into forensic practice is still nascent and continues to develop (Rawashdeh, 2025).

What is clear is that the field of auditing is on the move. Sathe (2025) anticipates that the incorporation of AI in forensic accounting will mark a turning point in the way fraud is detected and investigated, given that tools based on these models make it possible to identify anomalies, recognize suspicious patterns, and streamline investigation processes. Ellili et al. (2024) also note a sustained growth in publications on the subject, which accounts for an academic community that is taking this disciplinary crossover seriously and add that digital disruption, driven by technologies such as the cloud, social networks and massive data analysis, generates unprecedented volumes of structured, semi-structured and unstructured information that open up opportunities and challenges for fraud detection.

However, the accounting literature also reveals important gaps. A. Y. A. B. Ahmad (2024) and Moll and Yigitbasioglu (2019) agree on the need to delve into the implications of AI for the future of the accounting profession, especially in terms of roles, competencies, and ethics. Other studies warn that the widespread adoption of AI-based accounting systems remains constrained by users' distrust of their reliability, transparency, and governance (Dwivedi et al., 2021). A third group of studies has addressed the impact on the human factor in a fragmented way (Damioli et al., 2021; Frank et al., 2019) or the technological dimension (T. Ahmad et al., 2021; Ahuja & Nair, 2021; Briganti & Le Moine, 2020), without sufficiently integrating both perspectives in the forensic field.

This article seeks to fill some of that gap. Based on a systematic review of the literature of the last decade in Scopus, it was analyzed how AI and machine learning are being used in forensic accounting for fraud detection, with particular attention to the tools employed, the most frequent methodological approaches and the empirical evidence on their effectiveness.

The central question guiding the study is: To what extent do institutional, contextual, and information quality conditions determine the effectiveness of artificial intelligence and machine learning as fraud detection tools in forensic accounting, based on empirical evidence from the last decade?

The article is structured as follows: first, it presents the state of the art on forensic accounting, AI and machine learning applied to fraud detection; then it describes the PRISMA methodology used for the selection and analysis of studies; the results are then presented and discussed, highlighting trends, opportunities and challenges; and it ends with the conclusions, practical implications and recommendations for future research.

2. Literature Review

Before examining how artificial intelligence is integrated into forensic accounting, it is necessary to situate an analysis in the conceptual framework that defines what fraud is and why it occurs. The Fraud Triangle, formulated by [Cressey \(1953\)](#) based on his research on embezzlers, establishes that fraud requires the coexistence of three conditions: pressure, understood as the incentive or motivation that leads an individual to act dishonestly; timeliness, referring to weaknesses in the control systems that make it possible to commit the act without being detected; and rationalization, that is, the cognitive justification by which the individual normalizes his fraudulent behavior. [Wolfe and Hermanson \(2004\)](#) expanded this model with the Fraud Diamond, adding a fourth dimension: capacity, understood as the technical skill and positional access that allow the individual not only to want to commit fraud but to execute it effectively.

These theoretical perspectives on the causes of fraud are complemented by the practical frameworks developed by the foundational forensic accounting literature. [Albrecht et al. \(2012\)](#) established that forensic investigation requires the systematic integration of accounting techniques, interviewing skills, and document analysis to construct valid evidence, stressing that no technology can replace the judgment of the expert investigator but only extend his analytical capacity. [Singleton and Singleton \(2010\)](#), on the other hand, developed a comprehensive framework for fraud investigation that articulates detection procedures with the requirements of evidentiary admissibility in court, establishing that the forensic utility of any finding depends not only on its technical accuracy but also on its documentary traceability and its ability to hold up to legal scrutiny. Both frameworks anticipated, from forensic practice, the distinction that the recent AI literature is beginning to articulate between detecting an anomaly and producing legitimate forensic evidence.

These classic frameworks define the problem that forensic accounting and, more recently, artificial intelligence are trying to solve. Machine learning models act primarily on the opportunity dimension of the Fraud Triangle by continuously processing large volumes of transactions, reducing the time window in which an irregularity can remain undetected, and closing information gaps that manual control systems cannot fill ([Ramzan & Lokanan, 2025a](#)). Natural language processing techniques provide additional capability when analyzing unstructured text such as notes to financial statements, internal communications, and management reports, where linguistic cues associated with the rationalization dimension can be identified, such as minimizing risk language or the absence of explanations for significant variations. Models that incorporate behavioral and network variables, such as those developed by [Baghdasaryan et al. \(2022\)](#) and [Zhou et al. \(2023\)](#), can approximate the pressure dimension by identifying relational and contextual patterns that statistically precede fraudulent behavior.

This articulation between classical fraud theory and the technical capabilities of AI indicates that the most useful models for forensic accounting are not those that maximize classification accuracy in historical data, but those that are designed to detect the structural

conditions that make fraud possible, i.e., those that operate on opportunity, rationalization and capacity before pressure is translated into action. This perspective guides both the selection of relevant variables and the interpretation of the study's findings.

In relation to artificial intelligence and forensic accounting, it has become an area of sustained academic production, with tools that no longer only process data but also learn from it. Understanding this journey requires reviewing both the conceptual foundations that support it and the empirical findings that nourish it (Qatawneh, 2024). An inevitable starting point is the realization that financial systems generate volumes of information that exceed the capacity of manual analysis. Faced with this problem, machine learning techniques offer a concrete answer and make it possible to examine large sets of transactions continuously, locate statistically unusual patterns and translate that search into useful red flags for the forensic investigator (Esmeralda et al., 2026). Roszkowska (2021), whose work on fintech in financial reporting remains the most cited in the field with 162 cumulative references, clearly argued that blockchain, the Internet of Things, and AI models are not only tools for operational efficiency but also a redesign of the architecture on which forensic accounting and fraud prevention operate.

In that order, Khan et al. (2022), in a study of listed companies in the United States that accumulates 76 citations, combined the search algorithm for beetle antennas with supervised classification techniques to detect fraud in financial statements. The result was superior to previous reference models in computational time consumption, which raised a question that other authors would continue to explore, and the efficiency of a model is not only a matter of accuracy, but also of speed and responsiveness in environments where fraud evolves.

That same tension between accuracy and adaptability appears in Baghdasaryan et al. (2022), who worked with Armenian fiscal data and demonstrated that gradient boosting models significantly improve rule-based approaches when network variables, the links between suppliers and buyers, are incorporated as predictors. With 42 citations, their study is especially relevant to the present work because it shows that the performance of AI models does not depend only on the algorithm chosen but on the quality and richness of the data with which it is trained.

In parallel, the literature has paid attention to who operates these tools. Khaksar et al. (2022) analyzed the relationship between auditor characteristics such as experience, turnover, sector specialization, and effectiveness in fraud detection. Their work, with 36 citations, suggests that AI does not operate in a vacuum, since its usefulness depends to a large extent on the profile of the professional who interprets its results and the institutional trust that he or she places in automated systems. This observation connects directly with the warnings of Sathe (2025), who points out that, although anomaly detection algorithms can process financial information with a speed and coverage impossible for manual analysis, their real value is activated when the forensic accountant knows how to read and question them.

The work of Chen and Wu (2023), with 35 citations and focused on companies listed in China, offers one of the most comprehensive comparative analyses between individual algorithms and assembly methods for fraud detection in financial reports. Their central finding, indicating that assembly algorithms, and in particular stacking, consistently outperform individual models, reinforces the idea that the integration of multiple analytical perspectives produces more robust results than any single technique. This finding has direct practical implications for forensic teams, who can combine models with different biases and strengths to reduce false negatives.

Apart from the most cited articles, the recent literature adds important nuances. Akinbowale et al. (2023) studied the integration of forensic accounting and big data in South African banking, demonstrating that neural networks with hidden layers can classify

indicators of internal fraud with 95% correctness, while warning that continuous updating of models is as important as their initial design. In a complementary direction, [Zhou et al. \(2023\)](#) proposed FraudAuditor, a visual analytics system that integrates community detection with scanning interfaces for collusive fraud in health insurance. His proposal shows that the most useful AI in forensic contexts is not always the most sophisticated, but the one that can be validated and interpreted by the human analyst.

This need for interpretability runs through much of the recent literature. [Gkegkas et al. \(2025\)](#) systematized 43 empirical and theoretical studies on data analytics in the detection of fraud in financial statements. They identified a persistent gap: the lack of explainable models that can be sustained before regulatory bodies. [Tragouda et al. \(2024\)](#), on the other hand, proposed a multi-label classification approach for Greek financial statements that not only detects whether there are irregularities but also distinguishes their nature, a forensic distinction of the first order when it comes to constructing valid evidence in court.

The technological dimension expands when blockchain is incorporated, [Roszkowska \(2021\)](#) already anticipated this convergence, and subsequent work such as that of [Zandi et al. \(2025\)](#) developed it in more detail, where the combination of distributed and immutable records with predictive machine learning models creates an environment where the quality of the data improves, because it is more difficult to manipulate, while models have more to work with. This synergy is especially relevant for forensic accounting, where the traceability of transactions is not a technical detail but the core of any expertise.

The sectoral dimension also deserves attention. [Fourkiotis and Tsadiras \(2025\)](#) applied a complete AI pipeline, including variable selection with Boruta, categorical coding and threshold calibration by fold, to hospital claims data, demonstrating that discipline in preprocessing matters as much as the choice of classifier. [Murorunkwere et al. \(2022\)](#) did this in the Rwandan tax field, using artificial neural networks to identify factors of income tax fraud with an accuracy of 92%. Both studies illustrate that the techniques developed in advanced financial contexts are transferable to very different institutional environments, provided that the variables are adapted and the representativeness of the training data is taken care of.

Natural language processing opens another avenue. [Qatawneh \(2024\)](#) studied how NLP moderates the relationship between AI and fraud detection in accounting information systems, concluding that its moderating effect is statistically significant, where models that incorporate unstructured text analysis, notes to financial statements, internal communications, and management reports detect anomalies that numerical models miss. [Kurniawan et al. \(2026\)](#) corroborated this finding in Indonesia, where the analysis of annual reports using TF-IDF and logistic regression outperformed the rule-based manual audit method by 88.9%.

Finally, the literature on explainable AI (XAI) represents the most active frontier in the field. [Silahtaroglu et al. \(2026\)](#) integrated variable engineering with gradient boosting models and permutation significance analysis to predict the risk of modifications in the audit opinion, offering the auditor not only a prediction but also an interpretable justification for each finding. [Ramzan and Lokanan \(2025a\)](#), in a systematic review of 88 articles, showed that the literature is evolving from classical statistical methods to computational techniques, but that the articles that reach the journals with the highest impact are precisely those that show complete performance metrics and solid conceptual frameworks, not just high accuracy.

2.1. Theoretical Studies

The incorporation of artificial intelligence in forensic accounting does not occur in a conceptual vacuum. Behind every decision to adopt or resist a technology, to integrate it

into existing processes, or to deploy it in specific institutional contexts, there are theoretical assumptions about how organizations work, what legitimizes their practices, and what value the information they produce has. Three theoretical frameworks are mainly relevant to understanding why AI does not transform forensic accounting in a uniform or automatic way: contingency theory, institutional theory, and information value theory.

In contingency theory and contextual adaptation of models, [Lawrence and Lorsch \(1967\)](#) argued that there is no optimal and universal way to organize business management; what works in an environment depends on the particular conditions of that environment. Transferred to forensic accounting, this premise has direct consequences, since a machine learning model trained with data from companies listed in developed markets does not behave in the same way when applied to SMEs in emerging economies, financial cooperatives or public sector entities with heterogeneous information structures. The effectiveness of the system is not a property of the algorithm; it is the result of its adjustment to specific organizational and institutional conditions.

This argument finds empirical support in several works of the corpus analyzed. [Baghdasaryan et al. \(2022\)](#) clearly illustrated this when working with Armenian taxpayers; the variable that most improved the performance of their *gradient boosting* models was not technical, but contextual, in the network links between suppliers and buyers, information that only makes sense when the structure of the local market is known. Similarly, [Murorunkwere et al. \(2022\)](#) showed that the optimal neural network architecture for detecting tax fraud in Rwanda, without hidden layers and with a *softsign* activation function, differed from the usual configurations in the Western literature, precisely because the characteristics of the problem changed with the context. This evidence suggests that contingency is more than a philosophical framework; it is a practical guide since, before deploying any AI tool in forensic accounting, it is necessary to analyze the conditions of the environment, such as the quality of the available data, the structure of transactions, the capabilities of the audit team, and variables that condition the result, such as the model itself.

In institutional theory and the legitimacy of technological adoption, [Meyer and Rowan \(1977\)](#) argued that organizations not only adopt structures and practices because they are technically efficient, but because they confer legitimacy on them in their institutional environment. This distinction is important to understand why AI is advancing at such different speeds in different accounting and regulatory systems. It is not enough for a model to detect fraud more accurately; for its use to be consolidated, it must be recognized as valid by auditors, regulators, courts and supervisory bodies.

[Bracci et al. \(2026\)](#), in their study on supreme audit institutions, documented how isomorphic pressures, the need to appear modern and trustworthy to stakeholders, drive the adoption of data mining technologies even when significant internal resistance persists. Their analysis of 206 questionnaires showed that the strongest adoption inhibitors were not technical but organizational, such as anxiety about technology, perception of loss of professional control, and distrust in the system's results. [Aldemir and Uçma Uysal \(2024\)](#), on the other hand, identified that public sector internal auditors recognize the need for AI competencies but rarely develop them without an institutional framework that explicitly requires them. These findings confirm what Meyer and Rowan anticipated almost five decades ago: technological adoption responds to both legitimation logics and efficiency criteria, and any implementation strategy that ignores the institutional dimension is destined to face resistance that no algorithm can solve.

This perspective also illuminates the geographical inequality shown by the reviewed academic production. The leadership of the United States, China, and Australia in AI and forensic accounting research does not reflect only a technical advantage; it responds to regulatory systems that require greater accountability, more developed capital markets and

legal frameworks that give evidentiary value to the results of computational models. In contexts where this institutional scaffolding is more fragile, such as much of Latin America or Africa, adoption faces obstacles that go beyond the lack of technological infrastructure.

Regarding the theory of the value of information and the quality of data as a condition of possibility, [Arrow \(1971\)](#) proposed that information is an economic resource with strategic value, and that the information asymmetry between the parties to a transaction generates inefficiencies that can lead to market failure. In the context of forensic accounting, this idea takes on a technical dimension: if the data on which AI models are trained is incomplete, biased, or of low quality, the results they produce are not only inaccurate but actively misleading.

The recent literature has returned to this problem with increasing insistence, [Huang et al. \(2022\)](#) demonstrated working with fraud data in financial statements, that the treatment of missing values affects the performance of models, such as the choice of classifier, imputing with zero produced better results than the analysis of complete cases, but that conclusion depended on the type of model and the percentage of missing data. This technical finding has direct theoretical implications since the quality of the information is not a condition that can be taken for granted, but a variable that must be actively managed before any model comes into operation. [Chen and Wu \(2023\)](#) reinforce this point by pointing out that the superior performance of assembly algorithms over individual models is partly explained by the fact that they are more robust to heterogeneity and noise in the data, a feature that is especially valuable in environments where financial reporting is not standardized.

The convergence of these three theories produces a more complex and more realistic picture of what it means to integrate AI into forensic accounting. It is not a matter of installing superior technology that replaces obsolete procedures. It is about adjusting analytical tools to specific organizational contexts, such as contingency, building the institutional legitimacy necessary for their results to be accepted and used for institutional value, and ensuring that the information on which these tools operate is of sufficient quality to produce valid findings known as the value of information. Table 1 summarizes these articulations.

Table 1. Theories that support the application of AI in the detection of accounting fraud.

Theory	Author(s)	Main Focus	Application to AI, IAoT, and Accounting
Institutional Theory	Meyer and Rowan (1977)	Adoption of practices to gain legitimacy in the regulatory and social environment	It justifies the incorporation of AI to strengthen institutional credibility and respond to the accountability demands of supervisory bodies.
Contingency theory	Lawrence and Lorsch (1967)	Adaptation of practices according to the environment and organizational complexity	It explains why ML performance in fraud detection varies depending on the specific conditions of each organization, industry, and market.
Information Theory of Value	Arrow (1971)	Information as a strategic resource for efficient decisions	AI enhances forensic accounting by improving the quality, timeliness, and analysis of data, a necessary condition for detecting irregularities early and reliably.

Source: own elaboration.

The convergence of these three theoretical frameworks is fundamental in this research, since, initially, institutional theory operates as a central element of this study by explaining why the adoption of artificial intelligence in forensic accounting does not respond to an exclusively technical logic, but to legitimacy pressures that vary between regulatory systems, organizational cultures and jurisdictional contexts (Ramzan & Lokanan, 2025b). Within this framework, contingency theory fulfills a complementary function by specifying the conditions under which a machine learning model produces useful forensic results in an organizational context given the quality and representativeness of the available data, the structure of the transactions, the capabilities of the audit team and the characteristics of the local regulatory environment. The information value theory, on the other hand, identifies data quality as the variable that most often explains the gap between the theoretical performance of models and their practical effectiveness in real forensic environments. Articulated in this way, the three theories allow us to interpret the bibliometric patterns identified in this review as evidence of the state of institutional legitimation, contextual adaptation, and informational maturity that the field has reached in the last decade. This articulation guides both the reading of the results and the construction of the future research agenda.

The previous theoretical and foundational frameworks presented in this section guide the reading of the empirical evidence. Institutional theory explains why empirical studies show geographically asymmetrical patterns of adoption. Contingency theory explains why the performance of models varies according to the organizational and institutional context in which they are deployed. Information value theory explains why data quality is the variable that most often differentiates studies with robust results from those with persistent limitations. And the foundational frameworks of Cressey (1953), Wolfe and Hermanson (2004), Albrecht et al. (2012) and Singleton and Singleton (2010) establish the problem that all that empirical evidence tries to solve: fraud is not only a statistical anomaly but a human phenomenon with structural causes that technology can detect but not eliminate on its own. With that reading framework in place, the next section examines what the empirical evidence in the corpus says about how AI is responding to that challenge.

2.2. Empirical Studies

The recent literature offers a heterogeneous picture of the technical results and warnings of automation, where advances in developed economies contrast with an adoption that is still incipient in more fragile institutional contexts. For decades, fraud detection in forensic accounting depended fundamentally on the judgment of the auditor. This approach, although valid, faces a structural limitation that has become more acute with the digitization of organizations, because the number of transactions that an audit team must examine far exceeds what is possible to review manually with reasonable cost and time criteria. It was precisely that gap that drove academic and professional interest in machine learning models as a non-substitute complement to forensic work.

Rawashdeh (2025), who documented this transition from the perspective of accounting professionals themselves, through a survey distributed in Jordan, found that the adoption of AI is reforming decision-making processes in auditing and generating shifts in traditional roles, although with significant resistance among those who perceive automation as a threat to their expertise. This finding is relevant because it situates the problem in its human dimension, where it is evident that technology does not operate alone, but through professionals whose willingness to use it and whose ability to critically interpret it determine, to a large extent, its real effectiveness (Banța et al., 2022).

The most consistent empirical evidence comes from studies that compare the performance of different algorithms against specific fraud detection problems. Bao et al. (2022), in

their review of machine learning models applied to fraud prediction in financial statements of U.S.-listed companies, confirmed that computational approaches consistently outperform traditional statistical models in classification accuracy, especially when working with large datasets and high-dimensional variables. Their analysis established a methodological benchmark that subsequent studies have continued to refine.

In the same vein, [Lokanan and Sharma \(2025\)](#) showed that machine learning models can exploit large accounting datasets to predict fraud in financial statements with superior performance to rule-based approaches, and that their usefulness grows when combined with non-financial variables, such as internal communications, board characteristics, and replenishment history, which these traditional models ignore. The forensic auditor working with AI has a broader set of red flags than conventional ratio analysis offers.

In the tax field, [Baghdasaryan et al. \(2022\)](#) proved that gradient boosting models substantially improve their performance when they incorporate network data such as links between taxpayers, suppliers, and buyers, this as predictor variables, suggesting that the relational structure of transactions contains information about fraud that individual indicators do not capture. In health insurance, [Zhou et al. \(2023\)](#) developed *FraudAuditor*, a visual analytics system that combines community detection with interactive scanning interfaces to identify collusive fraud between patients and providers. This indicates an emerging trend in the literature when addressing the most useful systems in forensic practice, which are not necessarily the most sophisticated in statistical terms, but those that allow the human analyst to validate, question, and interpret the results of the model.

[Herrerros-Martínez et al. \(2025\)](#) applied a hybrid clustering and forest isolation methodology to real business procurement data, demonstrating that the combination of unsupervised methods with explainability tools such as SHAP produces more adaptable and better justified suspicious transaction lists than global detection approaches. For the forensic accountant, this matters since it is not only about finding anomalies, but also about being able to explain to a client, a regulator or a court why a certain transaction was marked as irregular.

[Bello and Olufemi \(2024\)](#) concluded that new technologies that are enhanced with the help of AI can offer innovative solutions; organizations that adopt these technologies will be better prepared to face the future. The authors highlighted the importance of AI in fraud detection and the need for continuous improvements in innovation and research in this field.

[Vitali and Giuliani \(2024\)](#) study the impact of robotic process automation (RPA) and AI on audit firms through semi-structured interviews with Big4 and non-Big4 in Italy, where the need for professionals to have new skills in analytics and information technology is evident, as digital technologies transform traditional information technology models by relying on flexible infrastructures such as cloud computing, that shift on-premises hardware and software towards remote and scalable services, thus redefining the competencies needed to manage and exploit these technological environments.

Empirical evidence not only records progress but also accurately documents the limits of automation in forensic contexts. The first and most recurrent is data quality. [Huang et al. \(2022\)](#) showed that the treatment of missing values affects model performance as much as the choice of classifier, and that there is no universally superior imputation strategy, as the optimal decision depends on the type of model, the percentage of missing data, and the distribution of variables. In environments where accounting records are incomplete, inconsistent, or have been deliberately altered, this limitation is relevant.

The second limit has to do with interpretability. Authors such as [Saifudin et al. \(2025\)](#), in their systematic review on AI in auditing, pointed out that one of the most persistent obstacles to the adoption of deep learning models in forensic contexts is opacity, which appears when an algorithm detects an anomaly but cannot explain on what criteria it did so. The finding loses probative value in any judicial or regulatory process. This gap between

predictive capacity and interpretability is not a minor technical problem; it is a structural limitation that the emerging literature on explainable AI (XAI) is trying to solve, although with still partial results.

The third limit is ethical and organizational, where the automation of accounting processes raises questions about professional responsibility that the literature begins to address more seriously. When a model flags a company as fraudulent and that signal triggers investigations with legal and reputational consequences, who is responsible for the error if the model was wrong? Qatawneh (2024) addressed this tension by showing that natural language processing moderates the relationship between AI and fraud detection, but that its effectiveness depends on the professionals who use it understanding its scope and being able to question it, not delegate the final decision to it.

A growing body of empirical studies examines not the performance of models, but the organizational and formative conditions that determine whether those models are used well or poorly. Imjai et al. (2024), with empirical evidence from Gen Z accounting students in Thailand, demonstrated that big data analytics capabilities are closely linked to the development of diagnostic and forensic skills, and that this linkage is strengthened when training programs explicitly incorporate data analysis geared towards fraud detection. Their finding inverts the usual logic, because it is not a question of teaching AI on the one hand and forensic accounting on the other, but of integrating both in the same training process focused on real problems.

Ellili et al. (2024), in their bibliometric analysis of trends in forensic accounting research, identified that scientific production has been growing strongly, but that there is still a gap between works that develop technical tools and those that examine their practical application in organizations with different levels of digital maturity. This gap is especially pronounced in emerging economies, where the data infrastructure, the availability of trained professionals, and the regulatory frameworks that give legal validity to the results of the models are still under construction. Lescano-Delgado (2023), in a review of Latin American studies on AI and fraud control, found that most of the documented implementations are concentrated in the banking and financial sector, while applications in public auditing, the cooperative sector, and small businesses remain marginal.

Dimitropoulos and Reading (2025) explored how forensic accounting helps investigate high-level and state financial crimes through an interpretive qualitative approach, with relevant and impactful business cases and identified challenges such as unequal rules between countries and proposed greater global collaboration along with advanced techniques. In coherence, studies on the training of accountants in the age of data highlight that traditional accounting skills are not enough, but specific skills in forensic analytics and accounting are required.

The convergence of these findings forms a coherent empirical portrait where AI and machine learning substantially improve the ability to detect fraud in forensic accounting when quality data is available, professionals are trained to interpret its results, and institutional frameworks that recognize its value. When any of these conditions are missing, the technical promise is only partially fulfilled or not fulfilled at all.

3. Materials and Methods

3.1. Identification of Sources of Information

This article examined how artificial intelligence and machine learning are integrated into forensic accounting as relevant tools for accounting fraud detection. To this end, a systematic review of the literature was carried out using the PRISMA method, in order to ensure a rigorous, transparent and objective search and selection process. The search was conducted in the Scopus database, using a strategy built from key terms related to

the adoption of AI, the need for more efficient tools to identify anomalies in accounting data, and the challenges posed by their implementation in modern accounting practice. Figure 1 identifies selected studies that met previously defined inclusion and exclusion criteria, and their quality was assessed using a checklist. The review was conducted in accordance with the PRISMA protocol, and the PRISMA 2020 checklist is provided as Supplementary Material.

SYSTEMATIC REVIEW FLOW DIAGRAM (PRISMA 2020)

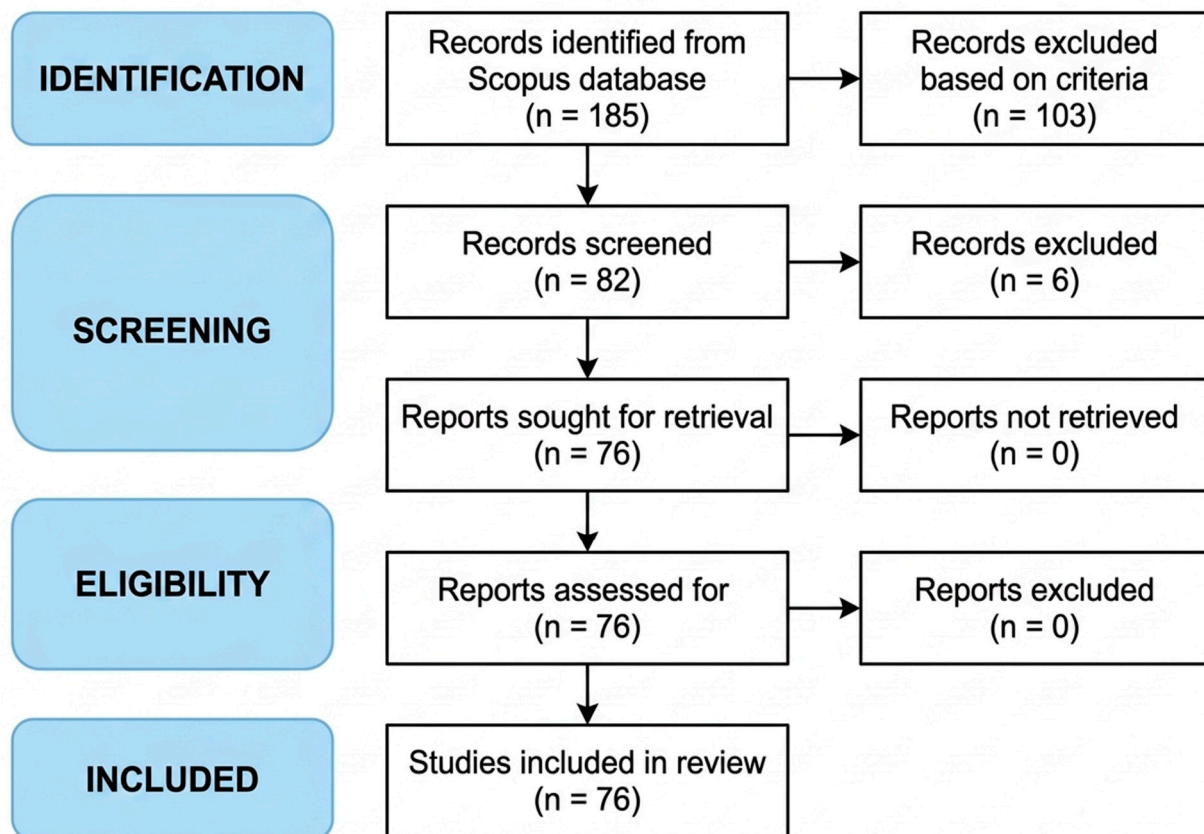


Figure 1. Prisma flowchart. Source: own elaboration.

The quality checklist was applied to all 76 documents after sifting abstracts and keywords. The criteria evaluated were the following: (1) clarity and explicitness of the research objective, (2) adequacy of the methodology to the stated objective, (3) transparency in the description of data sources and analytical procedures, (4) coherence between findings and conclusions, and (5) relevance of the study for the intersection between artificial intelligence, machine learning and forensic accounting or fraud detection. Each paper was independently evaluated by two members of the research team, and discrepancies were resolved through discussion and consensus. Studies that presented less methodological transparency were cited for their thematic relevance but were not used as primary evidence for the study's central analytical arguments. From these works, data were extracted on objectives, methodology, main findings and contributions to the field of study, which were later synthesized qualitatively with the help of Vosviewer and Bibliometrix de R to highlight emerging trends and research gaps.

3.2. Search and Data Extraction Strategy

A search strategy was designed using Boolean operators to identify articles in Scopus that jointly addressed audit, forensic accounting, machine learning, and artificial intelligence. Subsequently, the following search formula was constructed:

The extensive and systematic search in the Scopus database allowed the identification of 185 potentially relevant articles. To narrow down the analysis, the observation window was limited to the period 2019–2026 in the areas of economics, econometrics, finance, social sciences, business, and accounting. In addition, only articles and reviews in final versions in English or Spanish were selected (Table 2). By applying the inclusion and exclusion criteria (Table 3), the study base was reduced to 82 articles. Subsequently, the review of abstracts and keywords discarded 6 documents; consequently, the final synthesis included 76 manuscripts.

Table 2. Scopus search equation.

TITLE-ABS-KEY ("Forensic accounting" OR "audit" AND "accounting fraud" OR "financial fraud" OR "fraud detection" AND "artificial intelligence" OR "machine learning" OR "developing economies" OR "emerging markets") AND PUBYEAR > 2018 AND PUBYEAR < 2027 AND (LIMIT-TO (SUBJAREA, "COMP") OR LIMIT-TO (SUBJAREA, "BUSI") OR LIMIT-TO (SUBJAREA, "ECON") OR LIMIT-TO (SUBJAREA, "SOCI")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re")) AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "SPANISH")) AND (LIMIT-TO (PUBSTAGE, "final"))

Source: own elaboration.

It is important to note that the data corresponding to 2026 represents a year of partial indexing in Scopus at the time of running the search, so its volume of publications is naturally lower than in full years such as 2024 and 2025. This data point should not be interpreted as a slowdown of the field, but as an artifact of the cut-off moment of the search. The annual growth trends reported in this study should be read considering that the period 2019–2025 offers complete data, while 2026 reflects only the publications indexed as final versions up to the date of execution of the search in April 2026.

Table 3 contains the different inclusion and exclusion criteria used in the study.

Table 3. Inclusion and exclusion criteria.

Criteria	Inclusion	Exclusion
Publication date	Studies published between 2019 and 2026.	Articles that do not focus on machine learning or artificial intelligence were discarded.
Academic areas	Articles that address accounting aspects related to artificial intelligence and machine learning.	Studies that do not deal with accounting or artificial intelligence.
Language	Documents written in English and Spanish.	Languages other than English and Spanish.
Document Type	Research articles and reviews published in peer-reviewed academic journals.	Documents such as technical reports, theses, and book chapters.

Source: own elaboration.

The logic of each component of the search equation responds to deliberate methodological decisions described below to facilitate independent evaluation. The terms “Forensic accounting”, “accounting fraud”, “financial fraud” and “fraud detection” guarantee the coverage of the forensic domain and its most documented manifestations in the literature. The term “audit” was included because a significant part of the literature on fraud detection using AI is published under this disciplinary framework without explicitly using the term “forensic accounting”. The terms “artificial intelligence” and “machine learning” capture the study’s core technological approaches. The inclusion of “developing economies” and “emerging markets” was intentional and responds to one of the substantive interests of the research, which is to document the state of academic production on AI and forensic accounting in contexts beyond the most developed financial systems, given that the regional gap is one of the findings that the study seeks to identify and discuss. The subsequent application of filters by subject area (business, economics, computer science, and social sciences), document type (articles and reviews), language (English and Spanish), and publication status (final versions only) ensured that the resulting corpus was academically rigorous and thematically relevant, reducing the initial 185 results to the 76 manuscripts that make up the final synthesis.

3.3. Analysis of Bibliographic Data

The analysis of bibliographic data from the 76 selected documents was carried out in two steps following the approach proposed by [Meneses Cerón et al. \(2024\)](#). The first step consists of a descriptive analysis in order to identify the basic characteristics of the field of knowledge through descriptive statistics and the identification of the evolution of the scientific literature on the impact of the integration of AI in the detection of financial fraud. In addition, academic productivity was examined by year, country, sources and topic. This descriptive analysis library Figures 2–8 was performed using the Bibliometrix package executed from RStudio (version 4.4.2) with the Biblioshiny. The second step addressed bibliometric analysis, including co-citation analysis, keyword synergy analysis, identification of the most cited articles, and new trends in the field, using the analysis tool VOSviewer version 1.6.20 and Claude Opus, version 4.7 of Anthropic.

Bibliometrix made it possible to generate descriptive indicators of the corpus, maps of collaboration between authors and thematic distribution by keywords. VOSviewer was used to construct the co-citation network and the thematic density map, the outputs of which formed the visual basis of the analyses presented in Figures 9–11. The results produced by both tools were interpreted in direct relation to the research question and the institutional theoretical framework that guides the study, so that the bibliometric indicators are evidence of patterns of legitimation, contextual adaptation, and informational maturity of the field.

Claude Opus, version 4.7 of Anthropic, was used as a support tool for the interpretation of the thematic clusters identified by VOSviewer and for the conceptual organization of the emergent patterns of the corpus. Its use was limited to second-level analytical functions, identification of connections between clusters, suggestion of interpretative categories and support in the synthesis of cross-sectional patterns. In no case were primary analytical decisions or the direct extraction of data from the corpus delegated to Claude Opus. All the outputs generated by the tool were systematically verified by contrasting with the original data extracted from Scopus, Bibliometrix and VOSviewer. This procedure of human validation of AI outputs is consistent with the central argument of the study that artificial intelligence produces real analytical value when operating under competent professional supervision. The application of this principle to the methodological process

of the study itself constitutes, in itself, a practical demonstration of the thesis that the manuscript defends.

4. Results

The results of the bibliometric analysis are presented below. In the initial stage, the performance of the literature in terms of statistical descriptors, scientific productivity per year, most productive and cited authors, the countries with the greatest contribution to the production of articles and the most important scientific journals in the field of research on the integration of artificial intelligence and machine learning in forensic accounting as a key tool for the detection of accounting fraud are detailed.

The descriptive statistics of this literature review (Figure 2) reveal an academic interest in the application of artificial intelligence and machine learning in forensic accounting for the detection of accounting fraud. The annual growth of 45.79% between 2019 and 2026 reflects a sustained academic interest. The corpus comprises 76 documents from 60 different sources, indicating that the field does not depend on a small group of authors or journals but is progressively diversifying. International collaboration is present in 15.79% of the papers, with an average of 3.32 co-authors per article, indicating that research is being built under collaborative work and from different disciplines.



Figure 2. Statistical descriptors. Source: authors' elaboration based on Scopus and Bibliometrix.

The thematic diversity of the corpus is reflected in the 281 keywords recorded, indicating that the field is not limited to a single perspective but encompasses technical, legal, financial, and ethical dimensions. The almost 8896 references cited and the average age of the documents of just 1.86 years confirm that this is a recent field but with a theoretical basis in consolidation. The average of 10 citations per article shows that the academic community is actively incorporating this production into subsequent research. Together, these indicators signal a turning point in accounting and auditing practice, in which the integration of advanced technologies has ceased to be an option and has become a necessary condition for the effective detection and prevention of fraud in increasingly complex environments.

Figure 3 presents the annual distribution of the studies included in the corpus, showing a sustained growth trend between 2019 and 2026. The concentration of publications in recent years suggests that research interest in this field has progressively intensified, although it also indicates that certain methodological approaches and contexts of application concentrate a significant proportion of production. According to Qatawneh (2024), the boom in AI research for forensic accounting since 2022 is due to the adoption of sophisticated analytical, machine learning, and generative AI technologies in audits and fraud detection activities. Rapid digitalization after the pandemic increased exposure to financial risks and cyberfraud, which has driven the need for more predictive and automated control mecha-

nisms. Likewise, the consolidation of Audit 4.0 strengthened academic and professional interest in accounting models based on data intelligence and continuous monitoring.

On the other hand, this dynamic of growth of the literature around emerging technologies and fraud detection observed in Figure 3 shows the need to expand both the number and the geographical and disciplinary diversity of studies in this area, in order to generate knowledge that is more transferable to different institutional and organizational contexts.

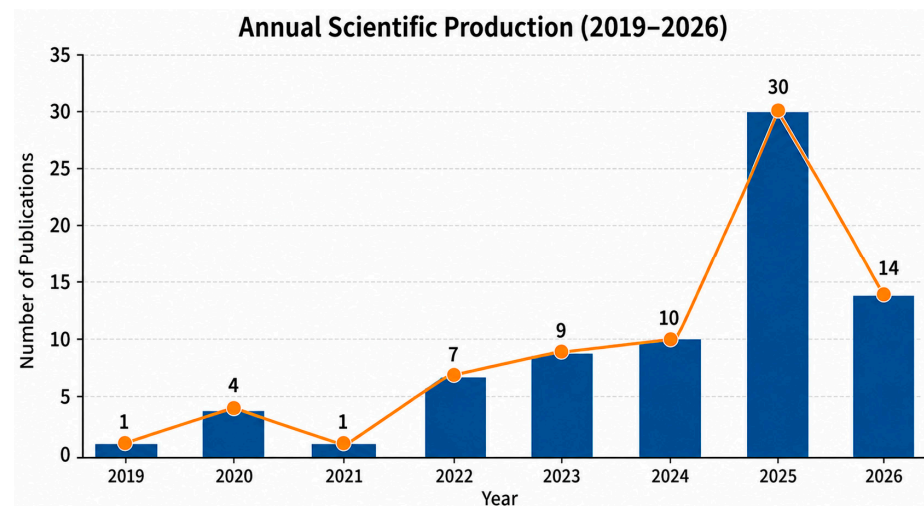


Figure 3. Articles published by year. Source: authors' elaboration based on Scopus.

Figure 4 shows the location of the most cited publications related to artificial intelligence, machine learning and their application in forensic accounting focused on accounting fraud. The United Kingdom stands out with a total of 239 citations, followed by China, Armenia and India, which have 132, 42 and 36 citations respectively. This concentration suggests a possible association between the contexts with the highest citation production and the existence of more demanding regulatory frameworks in terms of financial accountability, although bibliometric indicators alone do not allow establishing a causal relationship between the volume of citations and the level of institutional or regulatory development of each country.

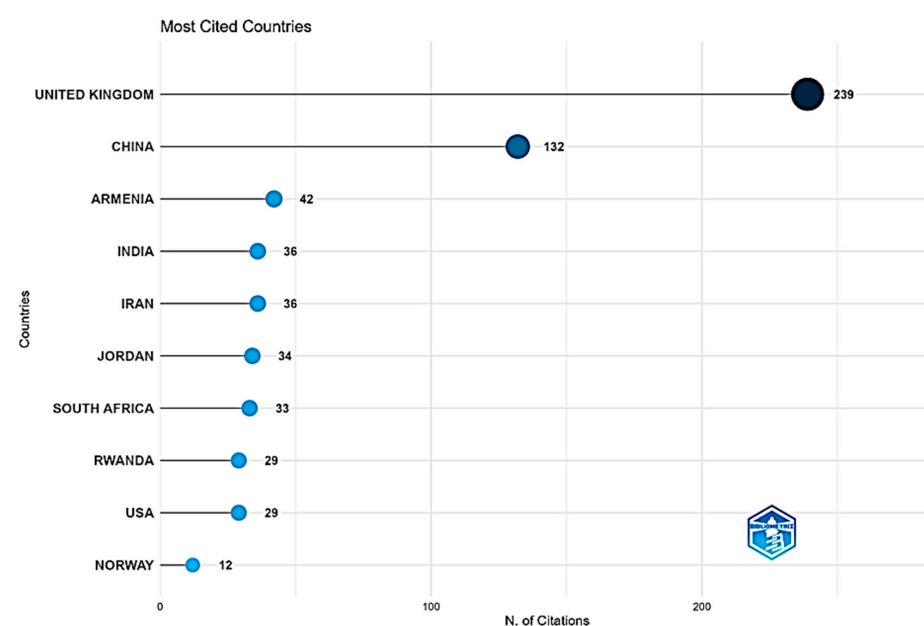


Figure 4. Articles published by country. Source: authors' elaboration based on Scopus and Bibliometrix.

From an accounting and financial perspective, this distribution indicates that economies with more developed regulatory systems are generating the methodologies and technologies that other jurisdictions could adopt to strengthen their anti-fraud mechanisms. At the same time, the lower participation of other regions in the production of knowledge represents both a challenge and an opportunity for the internationalization of forensic accounting research.

Figure 5 shows the distribution of the keywords in the analyzed corpus according to five major thematic areas, measured by frequency of occurrence. ML and AI methods lead with 44 occurrences, followed by fraud detection with 38, which together represent more than 74% of the total corpus frequency. This concentration reflects that the literature reviewed is fundamentally constructed from the technical and algorithmic dimension, where the development of predictive models and classifiers has concentrated the greatest research effort of the last decade.

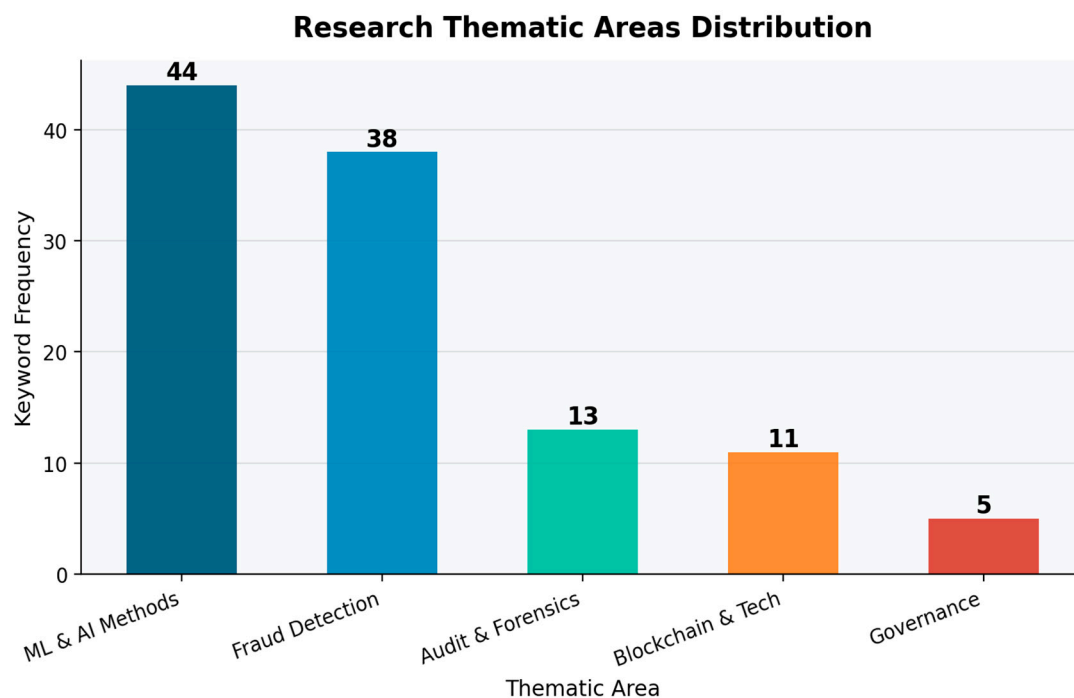


Figure 5. Articles published by subject area. Source: authors' elaboration based on Scopus.

Trailing far behind in the industry analysis, at a considerable distance appear auditing and forensic accounting with 13 occurrences, and blockchain and associated technologies with 11. The presence of these two areas in intermediate positions confirms a tension, while detection tools advance rapidly, specifically forensic frameworks, with their requirements of traceability, evidentiary validity and institutional adaptation, accumulate still significantly less production. Blockchain, on the other hand, appears as a technological component with a growing presence but is still subordinated to the dominant machine learning approaches. Finally, governance closes the list with only five occurrences, indicating that the institutional dimension of the field is underdeveloped compared to the technical one, which is especially striking, given that any effective implementation of AI in forensic accounting requires frameworks of responsibility, regulation and ethics that the technical literature tends to treat marginally. This low frequency confirms one of the most significant gaps in the field, where research is advancing faster in the design of algorithms than in the construction of the institutional and ethical frameworks necessary for these algorithms to operate legitimately in forensic and judicial contexts.

Figure 6 identifies the authors with the highest number of publications within the corpus analyzed. Three researchers stand out with 2 articles each: Chen Y. and Wu Z., co-authors of the study on the detection of financial fraud in listed companies in China using assembly algorithms, and Restrepo-Carmona J.A., a researcher at the National University of Colombia whose participation in two works on intelligent control systems reflects the growing Latin American interest in this field.

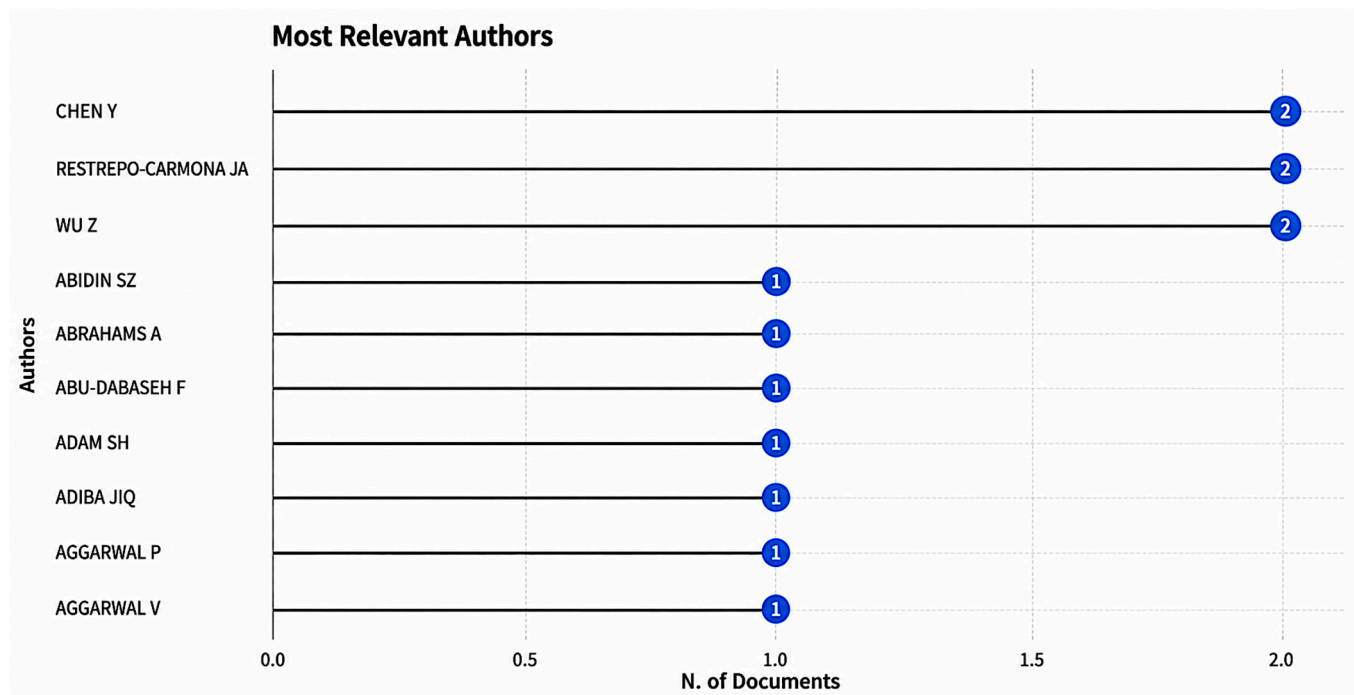


Figure 6. Most relevant authors. Source: authors' elaboration based on Scopus and Bibliometrix.

Similarly, Abidin S.Z., Abrahams A., Abu-Dabaseh F., Adam S.H., Adiba J.I.Q., Aggarwal P. and Aggarwal V, contribute one article each, which highlights a structural characteristic of the field: the academic production is highly distributed among individual authors who have not yet consolidated a continuous trajectory in this specific intersection between AI and forensic accounting.

This dispersion has a double reading. On the one hand, it indicates that the field continues to be permeable to new voices and disciplinary perspectives, which favors methodological diversity. On the other hand, the absence of authors with three or more publications suggests that consolidated individual references have not yet emerged to lead sustained research programs in this area. The relative exception is Chen Y. and Wu Z., whose continued collaboration on fraud detection in the Chinese market represents one of the few cases of progressive accumulation of knowledge within the reviewed corpus. This pattern reinforces the need to promote stable research teams and international collaborations that generate lines of work with greater depth and temporal continuity.

Figure 7 identifies the top academic journals that disclose research on artificial intelligence and machine learning in forensic accounting to detect fraud. *The Journal of Risk and Financial Management* of MDPI Publishing House, Switzerland, indexed in Scopus, ranked in Q2, leads with seven publications, followed by *Sustainability*, also by MDPI, indexed in Scopus and ranked in Q1 with three published papers. Eight journals, including *Intelligent Systems in Accounting and Management*, *IEEE Access* and *Future Internet*, make up the list, each with two publications. It is notorious that almost half of the sources are technical journals with technological approaches in the detection of accounting fraud. This

suggests that methodological advances are being driven more by artificial intelligence than by traditional accounting frameworks. This indicates, on the one hand, the imperative of greater interdisciplinary articulation and, on the other, the strengthening of publications in high-impact journals in accounting. The predominance of technical journals over accounting journals can be identified among the most productive sources. This confirms from institutional theory that the advances in the field are being legitimized mainly by the computational community, and that accounting and auditing institutions have not yet consolidated their role as producers of knowledge in this disciplinary intersection.

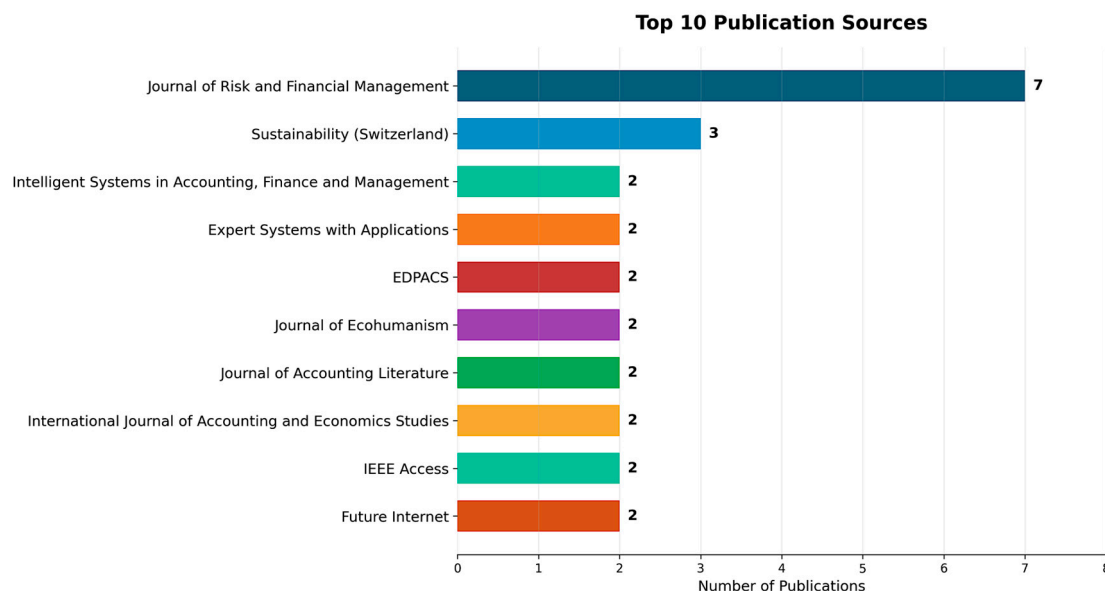


Figure 7. Articles published by the source. Source: authors' elaboration based on Scopus and Bibliometrix.

Figure 8 shows the institutional affiliations with the highest production of articles within the corpus analyzed. Diponegoro University in Indonesia tops the list with five publications, followed by Chitkara University in India with four. At a third level, with three articles each, appear American University of Armenia and Zhejiang University of China, which shows a geographically diverse participation that covers Central Asia, South Asia and the Arab world. With two publications each, Atlas SkillTech University (India), the Department of Mathematical Sciences and Global Studies affiliated with institutions in the United States and France, Islamic Azad University (Iran), Jain Deemed-to-be University (India), Jaipuria Institute of Management (India) and Jouf University (Saudi Arabia) complete the institutional map.

Several features of this distribution deserve attention. First, the strong presence of Asian institutions indicates that research on AI and forensic accounting is not an exclusive phenomenon of Anglo-Saxon economies; on the contrary, countries with fiscal and regulatory systems in transformation are generating relevant empirical contributions from their own contexts. Second, the absence of Latin American or African universities among the most productive institutions is consistent with the regional gap identified in the analysis by country, which may reflect both a lower research production in this area and limitations of indexing in Scopus, which has historically had less coverage of Latin American journals in accounting and computational sciences. Third, the fact that no institution exceeds five publications reflects the fragmentation characteristic of an expanding field, where knowledge is built from many fronts simultaneously without any center having yet reached a hegemonic position.

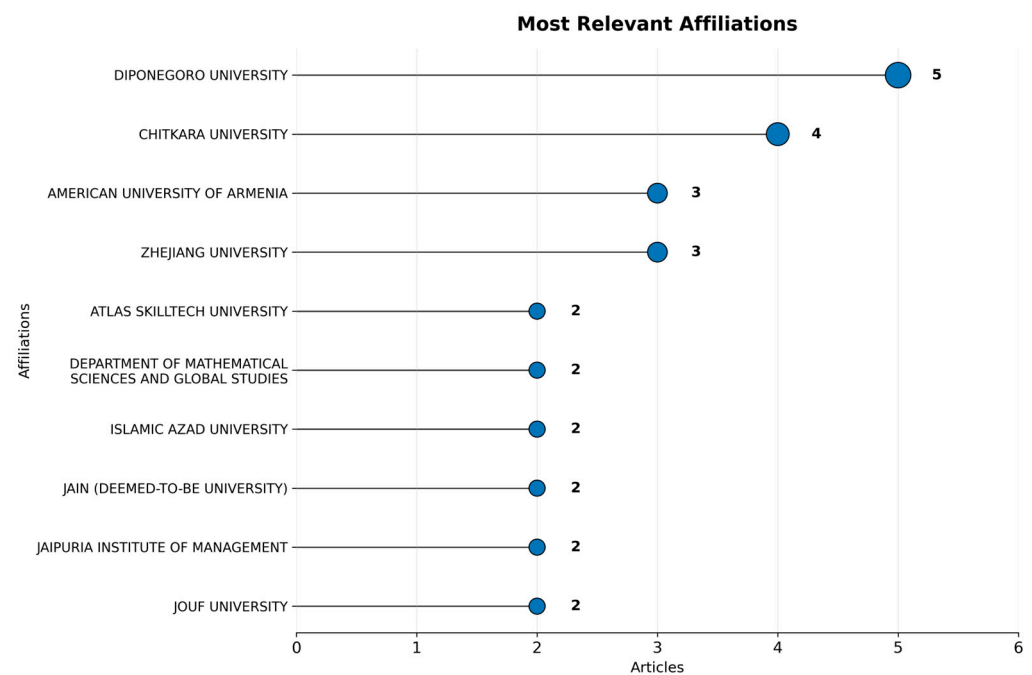


Figure 8. Articles published by study centers. Source: authors' elaboration based on Scopus.

Table 4 shows the five papers with the greatest citational impact within the corpus analyzed, all published between 2021 and 2023, confirming that the field is consolidating from a very recent bibliographic base. Roszkowska (2021) leads with 162 citations, a considerable distance from the second work, reflecting the foundational role that her analysis of fintech, auditing, and fraud prevention has played as a theoretical and methodological reference for subsequent studies. Its publication in the *Journal of Accounting and Organizational Change* is also significant: unlike the rest of the table, this work comes from an accounting journal, not a technique, suggesting that the bridge between technology and forensic practice was bridged earlier from the accounting side than from the computational side.

Table 4. Most cited documents.

Title	Authors	Year	Magazine	Citations
Fintech in financial reporting and audit for fraud prevention and safeguarding equity investments	Roszkowska P.	2021	<i>Journal of Accounting and Organizational Change</i>	162
Fraud detection in publicly traded U.S firms using Beetle Antennae Search: A machine learning approach	Khan A.T.; Cao X.; Li S.; Katsikis V.N.; Brajevic I.; Stanimirovic P.S.	2022	<i>Expert Systems with Applications</i>	76
Improving Tax Audit Efficiency Using Machine Learning: The Role of Taxpayer's Network Data in Fraud Detection	Baghdasaryan V.; Davtyan H.; Sarikyan A.; Navasardyan Z.	2022	<i>Applied Artificial Intelligence</i>	42
The relationship between auditor characteristics and fraud detection	Khaksar J.; Salehi M.; Lari DashtBayaz M.	2022	<i>Journal of Facilities Management</i>	36
Financial Fraud Detection of Listed Companies in China: A Machine Learning Approach	Chen Y.; Wu Z.	2023	<i>Sustainability (Switzerland)</i>	35

Source: authors' elaboration based on Scopus and Bibliometrix.

The remaining four papers, published in 2022 and 2023, share a common trait: they all apply machine learning algorithms to specific fraud detection or audit efficiency problems.

Khan et al. (2022) with 76 citations and Baghdasaryan et al. (2022) with 42 represent the most algorithmic approaches, metaheuristic optimization and gradient boosting respectively, while Khaksar et al. (2022) with 36 citations provide a less technical and more institutional perspective by examining how the characteristics of the auditor, experience, turnover, specialization, and condition detection capacity. Chen and Wu (2023), with 35 citations in just two years of circulation, show a speed of adoption that anticipates a growing impact on citations. In general, these five works illustrate the methodological diversity that characterizes the field: optimization approaches, assembly models, network analysis and studies on the human factor coexist, all oriented towards the same central problem.

Figure 9 offers an immediate visual representation of the concepts that structure the analyzed literature, where the size of each term reflects its frequency of occurrence in the corpus. Three terms visually dominate by far: machine learning, fraud detection and crime, which clearly confirms that the field is organized around a central problem, fraud as a criminal phenomenon and a family of technical solutions, machine learning, which the literature has adopted as the preferred response. However, it is highlighted that the field has not yet built a transversal legitimacy among disciplinary communities for the study of this research.



Figure 9. Keyword analysis. Source: authors' elaboration based on Scopus and VOSviewer.

At a second level of prominence appear artificial intelligence, machine learning, crime and learning systems, terms that expand the technical core and connect it with broader dimensions of the field. The simultaneous presence of machine learning and machine-learning as separate entries reflects a lack of terminological standardization that is characteristic of rapidly expanding fields, where different authors and journals have not yet converged on uniform writing conventions.

In the third level of frequency, terms stand out that reveal the sectoral contexts where the most research has been done: health care, health insurance and data mining point to the health industry as one of the most productive empirical scenarios, something consistent with the studies by Zhou et al. (2023) and Fourkiotis and Tsadiras (2025) reviewed in

this work. On the other hand, detection framework, detection method, deep learning, deep neural networks and prediction modeling confirm the weight of architectural and methodological approaches in the technical literature. The peripheral position of terms such as forensic accounting, governance approach, blockchain, accountability and corruption, which precisely represent the institutional, regulatory and forensic dimension of the field, is striking. This visual asymmetry between the technical center and the forensic periphery of the map clearly summarizes one of the most relevant findings of this review: research on AI applied to fraud detection has advanced considerably in the development of algorithmic tools, but it has not yet integrated conceptual frameworks with the same depth, evidentiary and institutional that define forensic work in the strict sense.

Figure 10 presents the collaboration network between the authors of the analyzed corpus, where each node represents a researcher, the size reflects its number of publications, and the color identifies the collaborative cluster to which it belongs. Links between nodes indicate direct co-authorship in at least one article in the corpus. The most immediate feature of the network is its total fragmentation; eleven independent clusters are identified, none of which establish links with the others. This complete absence of bridges between groups is the clearest sign that the field is still in a phase of decentralized construction, where teams work in parallel from their own institutional and geographical contexts without having yet woven transversal networks of collaboration.

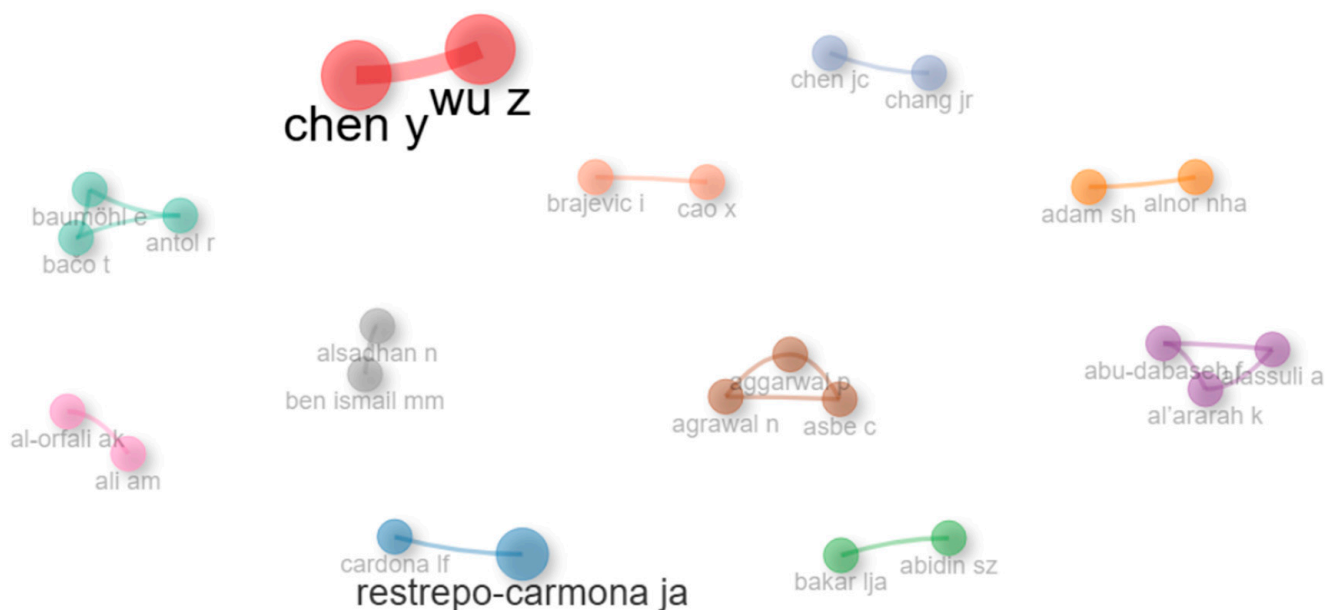


Figure 10. Collaboration network between authors. Source: authors' elaboration based on Scopus and Bibliometrix.

The most prominent cluster is formed by Chen Y. and Wu Z, the authors with the highest production of the corpus, whose sustained collaboration on fraud detection in Chinese companies makes them the only peers with a visible cumulative track record on the network. Restrepo-Carmona J.A. and Cardona L.F., researchers at the National University of Colombia, whose presence as the second largest cluster represents the most consistent Latin American contribution to the corpus, appear with similar weight.

The other clusters show structures of two or three nodes that correspond to point teams of co-authorship; Baumöhl E., Antol R. and Bačo T, linked to Central European universities; Brajevic I. and Cao X, from the team of Khan et al. (2022); Abu-Dabaseh F., Alassuli A. and Al'ararah K., originating from Jordanian universities; Aggarwal P., Agrawal

N., and Asbe C., from Indian institutions; and Adam S.H. and Alnor N.H.A., researchers from the Saudi and Sudanese fields.

It can be mainly determined that the geography of collaboration replicates the geography of production in countries such as Asia, the Middle East, Eastern Europe and Latin America, which appear as regional foci that investigate but autonomously without intersecting. Similarly, no author acts as a connector between clusters, which limits the circulation of methods and perspectives among research communities.

The total fragmentation of the network, with eleven clusters without any link between them, reflects, from the institutional perspective, the absence of mechanisms of transversal legitimation that encourage collaboration between disciplinary communities and regions. The field produces useful knowledge from parallel fronts, but it has not yet built the institutional bridges that would turn that dispersed production into an integrated scientific community.

Thematic Synthesis of the Corpus

The thematic analysis of the 76 documents of the corpus allows us to identify patterns of convergence and divergence that complement the bibliometric vision presented. In terms of AI and machine learning techniques, assembly and boosting methods, particularly XG-Boost and Random Forest, constitute the most frequent approach in the corpus, present in 20 studies, reflecting its consolidation as a methodological standard for the classification of fraudulent transactions in high-dimensional datasets. Data mining methods and statistical approaches appear in 22 studies, often as a complement to the supervised models. Deep learning and neural networks are present in 14 studies, with particularly relevant applications in the detection of bank fraud and in the classification of complex relational patterns. Explainable AI also appears in 14 studies, confirming a growing trend towards models that not only predict but also justify their findings, an essential condition for evidentiary validity in forensic contexts. Natural language processing and large-scale language models are present in five studies, representing the most recent frontier in the field.

As for forensic applications, audit quality assessment and audit risk management account for most of the corpus with 33 studies, followed by fraud detection in financial statements with 26 studies. Fraud in the public sector and corruption appear in 18 studies, while the literature review and bibliometric analyses add up to 14. Applications in banking and transactional fraud, tax fraud, and insurance and health fraud represent well-defined empirical niches with between four and seven studies each.

From the perspective of the institutional framework that guides this study, the thematic distribution of the corpus confirms that the field has advanced unevenly: the technical dimension, represented by the methods of assembly, deep learning and anomaly detection, concentrates the highest production density, while the institutional and governance dimension, which includes regulatory frameworks, the evidentiary validity of the models and the training of forensic professionals remains significantly underdeveloped. This asymmetry is a relevant finding of the review; the field increasingly knows how to detect fraud algorithmically, but has not yet resolved under what institutional conditions that detection produces legitimate and reliable forensic evidence.

5. Discussion and Future Research Agenda

The findings of this systematic review are interpreted through the institutional framework that guides the study. What the literature of the last decade shows is that the effectiveness of any artificial intelligence tool in forensic contexts depends on conditions that operate interdependently. These conditions are initially the institutional legitimacy of the environment in which it is deployed, in turn, the contextual adaptation of the model to

specific organizational characteristics, and finally, the quality of the information on which it operates. When any of these conditions fail, the technical promise is only partially fulfilled or not at all, regardless of the accuracy that the algorithm demonstrates under laboratory conditions. This conclusion, developed throughout the corpus analyzed, allows us to reorient the debate on AI and forensic accounting; the central problem is not technical but institutional, and its solution does not only involve better models but also better conditions for these models to operate with legitimacy, traceability and professional responsibility.

The findings of this review are consistent with what has been documented by [Bao et al. \(2022\)](#), [Lokanan and Sharma \(2025\)](#), and [Bello and Olufemi \(2024\)](#), who empirically show that machine learning models expand the detection capacity of the forensic professional without replacing their professional judgment. This distinction between computational detection and professional interpretation emerges recurrently in the corpus analyzed and constitutes, in the opinion of the authors, one of the most relevant findings for forensic practice.

Based on the evidence of the corpus, [Ramzan and Lokanan \(2025a\)](#) documented in their systematic review of 88 articles that the works with the greatest impact are not those that report the greatest algorithmic accuracy, but those that combine solid performance metrics with conceptual frameworks that explain the validity conditions of the models. This observation, derived directly from the reviewed literature, has implications that the authors consider relevant beyond the corpus, since it suggests that the development of more sophisticated models does not guarantee greater academic or practical contribution if it is not accompanied by a reflection on their institutional legitimacy and their usefulness for the forensic accountant.

At the level of vocational training, [Imjai et al. \(2024\)](#) empirically demonstrated that competencies in data analytics and forensic accounting are mutually reinforcing when training programs explicitly integrate them. This finding questions curricula that treat technology and accounting as separate modules; a forensic accountant who does not understand how an anomaly detection model works cannot critically interrogate it, and one who only knows how to program models without understanding the institutional context of fraud cannot interpret its results with evidentiary criteria. In this sense, [Figure 5](#) documents that governance is the concept with the least frequency in the corpus, with only five occurrences, compared to 44 for ML and AI methods and 38 for fraud detection. This asymmetry reveals that the research community has made considerable progress in the development of algorithms but has devoted marginal attention to the institutional, ethical and regulatory frameworks that determine whether these algorithms can be used in forensic and judicial contexts. The direct practical implication is that forensic auditor training cannot be limited to algorithmic training; it must integrate competencies in data governance, evidentiary validity of models, and professional responsibility when systems fail. [Imjai et al. \(2024\)](#), whose empirical evidence with Gen Z accounting students in Thailand shows that competencies in analytics and forensic accounting are mutually reinforcing when integrated into the same training process, confirm that this integration is operationally necessary for AI to produce real forensic value.

[Lokanan and Sharma \(2025\)](#) also point out that this professional updating faces real resistance in practice; a significant part of practicing accountants perceive automation as a threat to their expertise rather than as an extension of their capabilities. Overcoming that resistance requires not only ongoing training programs but also changes in the organizational culture of auditing firms and in the regulatory frameworks that define what counts as valid evidence in a forensic process.

5.1. Barriers and Opportunities in the Application of AI and ML in Forensic Accounting

The barriers to the adoption of AI in forensic accounting are not primarily technical, although they are often presented as such. The underlying problem is organizational and institutional; before any model comes into operation, it is necessary to ensure that the data with which it is trained is complete, representative and reliable. [Huang et al. \(2022\)](#) accurately showed that treating missing values affects model performance as much as algorithm choice, meaning that an organization with poor accounting records does not improve its forensics by simply installing an AI system; the model will learn the biases and gaps in the data it is fed and reproduce them in its predictions.

The second obstacle is interpretability. [Saifudin et al. \(2025\)](#) identified that one of the most consistent obstacles to the adoption of deep learning models in forensic accounting is their opacity: when the system detects an anomaly but cannot explain with what criteria it did so, the finding loses evidentiary value in any judicial or administrative process. This limitation, known in the technical literature as the black box problem, is not minor in a field where the traceability of evidence is a legal requirement, not a methodological preference. Advances in explainable AI (XAI) represent the most promising answer to this problem, as illustrated by [Silahataroğlu et al. \(2026\)](#), who integrated permutation significance analyses into their risk prediction models to provide interpretable justifications for each alert generated.

The third obstacle is regulatory in nature. [Remolina \(2025\)](#), in his analysis of Singapore's algorithmic governance framework applied to the financial sector, documented that even in jurisdictions with advanced regulatory frameworks, a gap persists between the available technical capabilities and the validation standards that regulators require to accept the results of the models as evidence. In emerging economies, where these frameworks are still under construction, this gap is greater.

The convergence between blockchain and machine learning, documented by [Zandi et al. \(2025\)](#) and conceptually anticipated by [Roszkowska \(2021\)](#), offers an architecture where data quality improves because records are more difficult to manipulate, while predictive models have more to work with. Figure 8 shows that no Latin American institution appears among the most productive in the corpus, and Figure 4 confirms that production is concentrated in the United Kingdom, China, and Armenia, contexts with more developed regulatory and data systems. These bibliometric patterns are consistent with the findings of [Huang et al. \(2022\)](#) in the analyzed corpus, who demonstrated that the treatment of missing values affects the performance of the models as much as the choice of classifier. The practical implication is institutional, since in economies with incomplete or heterogeneous accounting records, no investment in more sophisticated algorithms will produce reliable forensic results if it is not accompanied by an equivalent investment in data infrastructure, information quality standards and regulatory frameworks that require and verify that quality. Data governance is not a technical prerequisite but an institutional condition that, according to the information theory of value applied in this study, determines whether AI can operate with criteria of forensic validity and traceability. For forensic accounting, this means that the traceability of transactions, an essential condition of any expertise, can be guaranteed technologically rather than relying exclusively on document custody.

The integration of natural language processing with numerical models opens up another significant opportunity. [Qatawneh \(2024\)](#) demonstrated that systems that analyze both structured data and unstructured text, notes to financial statements, internal communications, and management reports, detect anomalies that purely quantitative models overlook, expanding the spectrum of signals available to the forensic investigator.

5.2. Future Research Agenda

The findings of this review clearly delineate several lines of work that the current literature has not developed in sufficient depth.

The first and most urgent is the development of explainable AI models specifically designed for forensic contexts. It is not enough for a model to be interpretable in statistical terms; their explanations must be understandable to auditors, judges, and regulators with varying levels of technical knowledge, and it must meet traceability standards that vary between jurisdictions. This line requires collaboration between researchers in computer science, law, and forensic accounting, a combination that this review's collaborative map shows is still exceptional.

The second priority line is research on federated learning applied to forensic accounting. This architecture allows multiple organizations to train fraud detection models collaboratively without sharing sensitive data with each other, simultaneously solving the privacy problem and the scarcity of labeled data that characterizes fraud sets. Its viability in small and medium-sized business contexts, where transaction volumes are lower and technological infrastructure is more limited, is an empirical question that the current literature has not answered.

The third line is longitudinal research on real effectiveness. Most of the studies reviewed evaluate their models on historical data through cross-validation, but they do not examine what happens when these models are deployed in real organizations for extended periods: how their performance degrades as fraud patterns evolve, what operational costs are generated to maintain and update them, and how the systematic generation of alerts that must be processed by human teams affects the organization. This gap between laboratory performance and impact in practice is one of the most relevant for professionals in the field.

The fourth line is research in emerging contexts. [Lescano-Delgado \(2023\)](#) documented that most AI implementations for fraud control in Latin America are concentrated in the banking sector, leaving public auditing, the cooperative sector, and small businesses unattended, which are precisely the contexts where vulnerability to fraud is greatest and institutional response capacity is most limited. Developing models adapted to the data, regulation, and infrastructure conditions of these economies is a task that the Latin American academic community is in a privileged position to lead, but that requires funding, data, and collaborative networks that do not yet exist with the necessary density.

Finally, the fifth line is the development of specific ethical and governance frameworks for forensic AI. [Figure 9](#) shows that terms such as forensic accounting, governance approach, and accountability occupy peripheral positions on the keyword map, while machine learning, fraud detection, and crime are concentrated in the center of the field. This distribution visually confirms that research is moving faster in building tools than in defining the conditions under which those tools are legitimate, responsible, and applicable in court. The most relevant practical implication that derives from this pattern is the development of specific ethical frameworks for forensic AI that answer three questions that the current literature has not answered in sufficient depth: Who responds when a model generates an erroneous alert with legal and reputational consequences? How are those investigations protected from algorithmic biases? And what auditing standards should be applied to AI systems themselves? [Remolina \(2025\)](#), whose analysis of Singapore's algorithmic governance framework documents that a gap persists between available technical capabilities and validation standards that regulators demand, illustrates that this is not a future problem but a present operational constraint.

Finally, the most relevant concepts of the literature related to artificial intelligence and machine learning are located in the context of forensic accounting and accounting

fraud detection according to their degree of development (density) and degree of relevance (centrality) (Figure 11). It is noted that issues such as “artificial intelligence”, “risk assessment”, “risk assessment”, and “investments” are in the quadrant of motor topics, which indicates that they are quite developed areas and quite connected to others, resulting in being pillars of the evolution of the field. Similarly, concepts such as “machine learning”, “deep learning”, and “decentralized finance” also figure as drivers, evidencing the rise in these topics in automated fraud detection. In the center of the map, topics such as “crime,” “detecting fraud,” and “machine learning” are placed as highly central, yet low-density, suggesting that they are basic topics which are fundamental, albeit informally structured. Conversely, terms such as “service industry” and “authentication” fall into the quadrant of emerging or declining issues, indicating subsectors that are growing or have no potential. All of this highlights the inadequacy of research and development at the interface of AI technologies with forensic accounting.

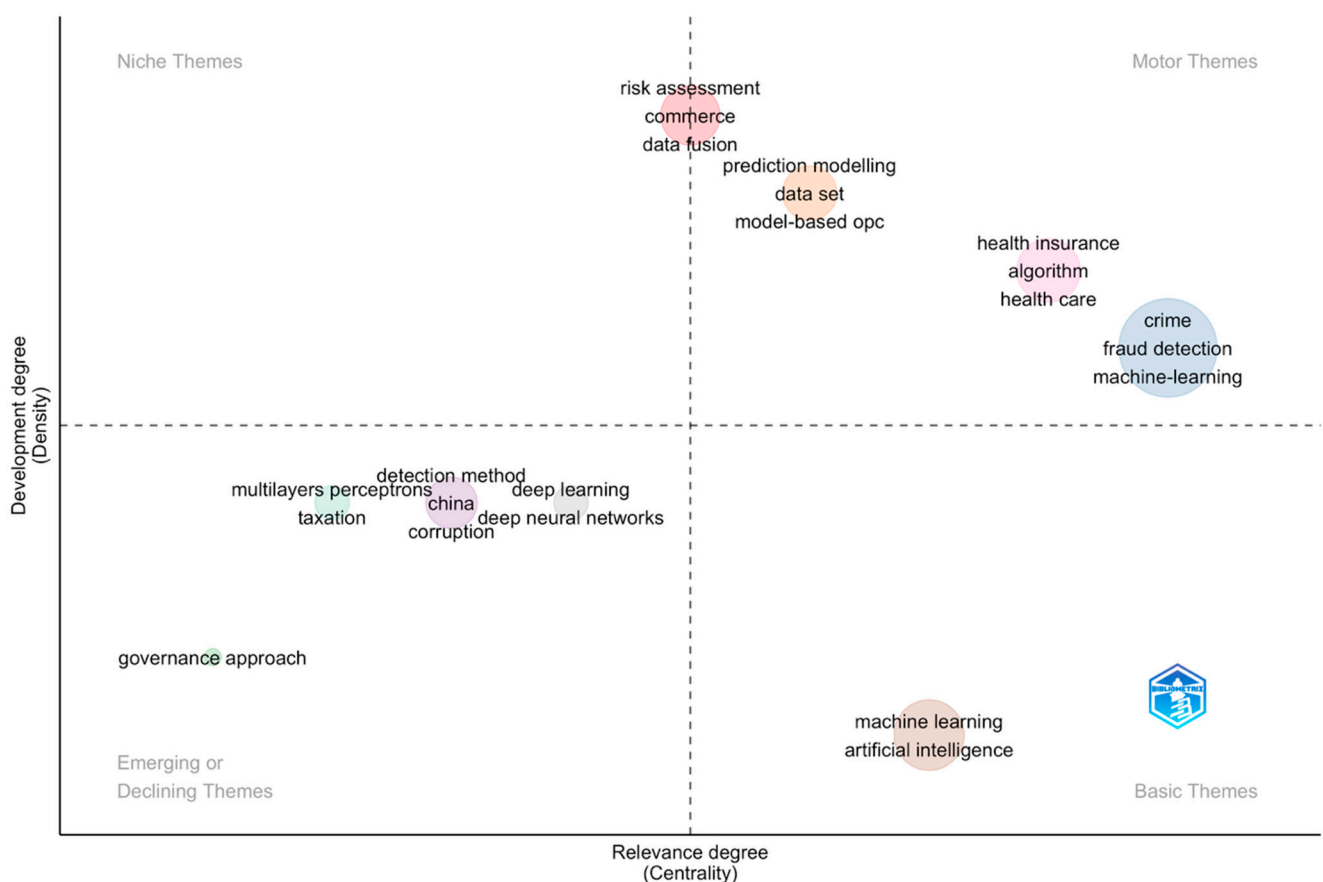


Figure 11. Future research agenda. Source: authors’ elaboration based on Scopus and Bibliometrix.

5.3. Limitations of the Research

The present study has limitations that should be considered when interpreting its findings. The first and most direct is the restriction to a single database such as Scopus, which guarantees coverage of high-impact indexed academic journals, but necessarily excludes a relevant part of the field’s scientific production. Since the intersection between artificial intelligence and forensic accounting is an inherently interdisciplinary area, significant studies may appear in conference proceedings of computer science and systems engineering, professional auditing literature and standards issued by bodies such as the IAASB or PCAOB, regulatory reports from financial supervisory entities, and specialized databases such as IEEE Xplore, ACM Digital Library or Web of Science. The exclusion of

these sources implies that the corpus analyzed mainly reflects academic accounting and management production and may underestimate the methodological advances that are published primarily in computational channels.

The second limitation concerns the construction of the search string. As explained in the methodological section, the inclusion of the terms “developing economies” and “emerging markets” in the Boolean equation responds to a substantive interest of the study, but may have introduced a selection bias by favoring articles that explicitly frame their analysis in emerging contexts, to the detriment of studies with equally relevant findings that do not use this terminology. Similarly, the inclusion of the term “audit” without subfield restriction may have incorporated general audit articles not directly related to forensic accounting. These decisions were deliberate and are justified in Section 3.2, but their consequences on the composition of the corpus must be taken into account when generalizing the findings.

The third limitation is the time window. The period 2019–2026 was selected to capture the most recent output in a rapidly evolving field, but excludes previous contributions that may have established relevant conceptual and methodological foundations, particularly seminal work on fraud detection published before the consolidation of deep learning as a dominant tool.

Finally, the interdisciplinary nature of the field introduces a structural limitation as computer science researchers, accounting professionals, and forensic accounting academics use distinct vocabularies to describe similar problems, making it inevitable that a search anchored in accounting terminology will underestimate output from technical communities, and vice versa. Acknowledging this limitation does not invalidate the study’s findings, but it does delimit its scope, since the corpus analyzed represents a systematic and transparent sample of academic production at the intersection of AI and forensic accounting indexed in Scopus, not an exhaustive coverage of all the knowledge available in the field.

6. Conclusions

The theoretical contribution of this study lies in having applied an integrated institutional framework to interpret the patterns that bibliometrics reveal, rather than simply describing them. By articulating institutional theory as the primary lens, complemented by contingency theory and information value theory, the study offers an explanation for why the adoption of artificial intelligence in forensic accounting is geographically asymmetric; moreover, organizationally uneven, and finally, technically promising but most importantly institutionally incomplete. This framework not only interprets the field’s past but precisely delineates what is urgent. Before developing more sophisticated models, the field needs to build regulatory frameworks that give evidentiary validity to its results, as well as training programs that allow auditors to interpret them critically, and governance standards that assign responsibility when systems fail. These elements are not of a technical nature but of an institutional nature.

The bibliometric analysis reveals a field in accelerated expansion, with academic productivity that has grown steadily between 2019 and 2026 and a publication base distributed across more than 70 different sources, indicating thematic diversity and disciplinary openness. The United Kingdom concentrates the largest volume of citations in the corpus, followed by China, Armenia and India. This asymmetrical geographical distribution, high citation in Anglo-Saxon economies, emerging production in Asia and the Arab world reflects that the field is globalizing, but that the most influential knowledge continues to be generated from contexts with more developed financial systems and more demanding regulatory frameworks.

The underrepresentation of Latin America in the corpus is a pattern documented in the bibliometric data of this review. Figure 4 shows the absence of Latin America among the regions with the highest production, and the findings of [Lescano-Delgado \(2023\)](#) in the corpus analyzed, who documented that most of the AI implementations for fraud control in the region are concentrated in the banking sector, leaving public auditing, the cooperative sector, and small businesses unattended. Consequently, this produces a practical involvement with a policy scope; the region is not free of vulnerability to fraud, but of applied research that develops tools adapted to its specific institutional conditions. Developing this research requires three conditions that currently do not exist with the necessary density: funding for access to quality data, collaboration networks between Latin American universities and international institutions, and legal frameworks that grant evidentiary validity to the results of the models in the region's judicial systems. Without these conditions, AI in forensic accounting will continue to be an imported phenomenon rather than a locally developed institutional capacity. Based on this corpus evidence, the authors argue that closing this gap is not only an academic task but a condition for these technologies to generate real value in the contexts where they are most needed, although they recognize that this statement goes beyond what bibliometric data alone can demonstrate and constitutes a normative interpretation based on the available evidence.

This interpretation is supported by the evidence of [Bracci et al. \(2026\)](#), who documented that the strongest inhibitors to AI adoption in auditing are not technical but organizational, and by the findings of [Aldemir and Uçma Uysal \(2024\)](#), who identified that public sector internal auditors recognize the need for AI competencies but rarely develop them without an institutional framework that explicitly requires them. These references to the corpus constitute the empirical basis of the statement; the normative implications that derive from them represent the interpretation of the authors.

One of the most revealing tensions that emerges from the analysis is the one between the technical capacity of the models and their institutional acceptance. The most statistically accurate AI systems are not necessarily the most useful in forensic practice if they cannot explain their findings in a way that auditors, regulators, and judges can validate. Explainable AI is therefore not an optional line of research: it is the sine qua non requirement for technical advances in the field to be translated into legitimate and reliable forensic tools.

The thematic map confirms this tension graphically; the concepts of greatest centrality and density in the corpus are fraud detection and machine learning, while forensic accounting, governance and explainable AI occupy peripheral or incipient development positions. In other words, the literature has advanced faster in constructing the tools than in defining the conditions under which those tools are valid, responsible, and applicable. This asymmetry precisely delimits the most urgent agenda of the field.

Finally, the study shows that the transformation of forensic accounting using AI is not a technical process but a social, institutional and training process. It requires auditors who understand the models they use, organizations that ensure the quality of their data, regulators that define validation standards, and ethical frameworks that assign responsibilities when systems fail. None of these elements are solved with a better algorithm. The most lasting contribution that research can make in this field is not to produce the model more accurately, but to build the knowledge necessary for existing models to be used well.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jrfm19070519/s1>, File S1: PRISMA 2020 checklist. [Page et al. \(2021\)](#) is cited in the Supplementary Materials.

Author Contributions: Conceptualization, L.Á.M.C., E.F.R. and V.E.P.T.; methodology, L.Á.M.C., E.F.R., R.C.V.S. and V.E.P.T.; formal analysis, L.Á.M.C. and E.F.R.; investigation, L.Á.M.C., E.F.R. and

R.C.V.S.; data curation, L.Á.M.C., E.F.R. and R.C.V.S.; writing—original draft preparation, L.Á.M.C., E.F.R., R.C.V.S., M.A.A.L. and J.J.F.S.; writing—review and editing, L.Á.M.C. and E.F.R. and V.E.P.T.; visualization, L.Á.M.C., R.C.V.S. and J.J.F.S.; supervision, L.Á.M.C., J.J.F.S. and M.A.A.L., funding acquisition: L.Á.M.C., and V.E.P.T. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded by the Corporación Universitaria Autónoma de Comfacaucá (UNICOMFACAUCÁ) and the Universidad del Cauca. Project number: 6657.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Ahmad, A. Y. A. B. (2024). Ethical implications of artificial intelligence in accounting: A framework for responsible AI adoption in multinational corporations in Jordan. *International Journal of Data and Network Science*, 8(1), 401–414. [\[CrossRef\]](#)
- Ahmad, T., Zhang, D., Huang, C., Zhang, H., Dai, N., Song, Y., & Chen, H. (2021). Artificial intelligence in sustainable energy industry: Status quo, challenges and opportunities. *Journal of Cleaner Production*, 289, 125834. [\[CrossRef\]](#)
- Ahuja, V., & Nair, L. V. (2021). Artificial intelligence and technology in the COVID era: A narrative review. *Journal of Anaesthesiology Clinical Pharmacology*, 37(1), 28–35. [\[CrossRef\]](#) [\[PubMed\]](#)
- Akinbowale, O. E., Mashigo, P., & Zerihun, M. F. (2023). The integration of forensic accounting and big data technology frameworks for internal fraud mitigation in the banking industry. *Cogent Business and Management*, 10(1), 2163560. [\[CrossRef\]](#)
- Albrecht, W. S., Albrecht, C. O., Albrecht, C. C., & Zimbelman, M. F. (2012). *Fraud examination* (4th ed.). Cengage Learning.
- Aldemir, C., & Uçma Uysal, T. (2024). AI competencies for internal auditors in the public sector. *EDPACS*, 69(1), 3–21. [\[CrossRef\]](#)
- Arrow, K. J. (1971). *Essays in the theory of risk-bearing*. Markham Publishing Company. Available online: <https://archive.org/details/essaysintheoryof0000arro> (accessed on 10 June 2026).
- Baghdasaryan, V., Davtyan, H., Sarikeyan, A., & Navasardyan, Z. (2022). Improving tax audit efficiency using machine learning: The role of taxpayer's network data in fraud detection. *Applied Artificial Intelligence*, 36(1), 2012002. [\[CrossRef\]](#)
- Banța, V.-C., Rindasu, S.-M., Tanasie, A., & Cojocaru, D. (2022). Artificial intelligence in the accounting of international businesses: A perception-based approach. *Sustainability*, 14(11), 6632. [\[CrossRef\]](#)
- Bao, Y., Hilary, G., & Ke, B. (2022). Artificial intelligence and fraud detection. In *Innovative technology at the interface of finance and operations* (Vol. 11, pp. 223–247). Springer. [\[CrossRef\]](#)
- Bello, O. A., & Olufemi, K. (2024). Artificial intelligence in fraud prevention: Exploring techniques and applications, challenges and opportunities. *Computer Science & IT Research Journal*, 5(6), 1505–1520. [\[CrossRef\]](#)
- Bracci, E., Tallaki, M., & Ebua Otia, J. (2026). Propensity factors of artificial intelligence technology adoption by public sector auditors. *Journal of Public Budgeting, Accounting and Financial Management*, 38(1), 59–86. [\[CrossRef\]](#)
- Briganti, G., & Le Moine, O. (2020). Artificial intelligence in medicine: Today and tomorrow. *Frontiers in Medicine*, 7, 27. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cappa, F., Oriani, R., Peruffo, E., & Scarpatetti, G. (2024). The power of big data affordances to reshape anti-fraud strategies. *Technological Forecasting and Social Change*, 213, 123456. [\[CrossRef\]](#)
- Chen, Y., & Wu, Z. (2023). Financial fraud detection of listed companies in China: A machine learning approach. *Sustainability*, 15(1), 105. [\[CrossRef\]](#)
- Cressey, D. R. (1953). *Other people's money: A study in the social psychology of embezzlement*. Free Press.
- Damioli, G., Van Roy, V., & Vertesy, D. (2021). The impact of artificial intelligence on labor productivity. *Eurasian Business Review*, 11, 1–25. [\[CrossRef\]](#)
- Dimitropoulos, G., & Reading, M. (2025). Forensic accounting as an investigative tool: Insights from the FTX and Qatargate. *Journal of Economic Criminology*, 7, 100132. [\[CrossRef\]](#)
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., & Eirug, A. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. [\[CrossRef\]](#)
- Ellili, N., Nobanee, H., Haddad, H., & Guellim, A. (2024). Emerging trends in forensic accounting research: Bridging research gaps and future directions. *Journal of Economic Criminology*, 4, 100065. [\[CrossRef\]](#)

- Erazo-Castillo, J., & De la A-Muñoz, S. (2023). Auditing of the future, foresight and artificial intelligence to anticipate risks in organizations. *Novasinergia*, 6(1), 105–119. [CrossRef]
- Esmeralda, A. I., Fadhilah, N. H. K., & Das, N. A. (2026). Fraud detection and machine learning in auditing: A systematic literature review. *Educational Studies and Research Journal*, 3(1), 37–46. [CrossRef]
- Fourkiotis, K. P., & Tsadiras, A. (2025). Future internet applications in healthcare: Big data-driven fraud detection with machine learning. *Future Internet*, 17(10), 460. [CrossRef]
- Frank, M. R., Author, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., & Moro, E. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14), 6531–6539. [CrossRef] [PubMed]
- Fu, B., Tong, Y., Li, Y., Tang, Z., Shang, Z., & Li, A. (2025). Financial fraud anomaly detection of listed companies based on probabilistic perspective machine learning models. *Computer Science*, 266, 979–986. [CrossRef]
- Gkegkas, M., Kydros, D., & Pazarskis, M. (2025). Using data analytics in financial statement fraud detection and prevention: A systematic review of methods, challenges, and future directions. *Journal of Risk and Financial Management*, 18(11), 598. [CrossRef]
- Herrerros-Martínez, A., Magdalena-Benedicto, R., Vila-Francés, J., Serrano-López, A. J., Pérez-Díaz, S., & Martínez-Herráiz, J. J. (2025). Applied machine learning to anomaly detection in enterprise purchase processes: A hybrid approach using clustering and isolation forest. *Information*, 16(3), 177. [CrossRef]
- Huang, L., Abrahams, A., & Ractham, P. (2022). Enhanced financial fraud detection using cost-sensitive cascade forest with missing value imputation. *Intelligent Systems in Accounting, Finance and Management*, 29(3), 133–155. [CrossRef]
- Imjai, N., Swatdikun, T., Rungruang, P., Basiruddin, R., & Aujirapongpan, S. (2024). Nurturing big data analytics, diagnostic, and forensic accounting skills for Generation Z accountants. *Journal of Innovation & Knowledge*, 9(3), 100123. [CrossRef]
- International Accounting Standards Board (IASB). (2018). *Conceptual framework for financial reporting*. IFRS Foundation. Available online: <https://www.ifrs.org/projects/completed-projects/2018/conceptual-framework> (accessed on 5 June 2026).
- Kanchan, M. (2021). The role of forensic accounting in fraud investigation. *International Journal of Research in Business Studies*, 6(2), 59–66. Available online: <https://www.ijrbs.com/wp-content/uploads/2021/12/Kanchan.pdf> (accessed on 10 June 2026).
- Khaksar, J., Salehi, M., & Lari DashtBayaz, M. (2022). The relationship between auditor characteristics and fraud detection. *Journal of Facilities Management*, 20(1), 79–101. [CrossRef]
- Khan, A. T., Cao, X., Li, S., Katsikis, V. N., Brajevic, I., & Stanimirovic, P. S. (2022). Fraud detection in publicly traded U.S firms using Beetle Antennae Search: A machine learning approach. *Expert Systems with Applications*, 191, 116148. [CrossRef]
- Kurniawan, A., Yudha, L. N., & Witjaksono, A. (2026). Improving accuracy in fraud detection in public companies financial reports using natural language processing. *Journal of Theoretical and Applied Information Technology*, 104(2), 192–204.
- Lawrence, P. R., & Lorsch, J. W. (1967). *Organization and environment: Managing differentiation and integration*. Harvard Business School Press.
- Lehner, O. M., & Knoll, C. (Eds.). (2022). *Artificial intelligence in accounting: Organisational and ethical implications*. Routledge. [CrossRef]
- Lescano-Delgado, M. (2023). Advances in the use of artificial intelligence to improve control and fraud detection in organizations. *Scientific Journal of Systems and Informatics*, 3(1), e494. [CrossRef]
- Lokanan, M., & Sharma, S. (2025). Reprint of: The use of machine learning algorithms to predict financial statement fraud. *The British Accounting Review*, 57(1), 101560. [CrossRef]
- Meneses Cerón, L. A., Ortega, F. J., Arcos Rodríguez, F. A., & Ramírez Mora, K. N. (2024). Accounting ethics as a pillar of corporate sustainability in emerging economies: A systematic review of the literature. *Revista En-Contexto*, 12(21), 257–286. [CrossRef]
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363. [CrossRef] [PubMed]
- Moll, J., & Yigitbasioglu, O. (2019). The role of internet-related technologies in shaping the work of accountants: New directions for accounting research. *The British Accounting Review*, 51(6), 100833. [CrossRef]
- Murorunkwere, B. F., Tuyishimire, O., Houghton, D., & Nzabanita, J. (2022). Fraud detection using neural networks: A case study of income tax. *Future Internet*, 14(6), 168. [CrossRef]
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. [CrossRef]
- Qatawneh, A. M. (2024). The role of artificial intelligence in auditing and fraud detection in accounting information systems: Moderating role of natural language processing. *International Journal of Organizational Analysis*, 33(6), 1391–1409. [CrossRef]
- Ramzan, S., & Lokanan, M. (2025a). Integrating criminological theories in accounting and finance fraud research: A systematic literature review. *Journal of Economic Criminology*, 9, 100179. [CrossRef]
- Ramzan, S., & Lokanan, M. (2025b). The application of machine learning to study fraud in the accounting literature. *Journal of Accounting Literature*, 47(3), 570–596. [CrossRef]

- Rawashdeh, A. (2025). The consequences of artificial intelligence: An investigation into the impact of AI on job displacement in accounting. *Journal of Science and Technology Policy Management*, 16(3), 506–535. [\[CrossRef\]](#)
- Remolina, N. (2025). AI governance and algorithmic auditing in financial institutions: Lessons from Singapore. *Journal of Risk Management in Financial Institutions*, 18(3), 261–275. [\[CrossRef\]](#)
- Renzhou, D. (2011). Research on legal procedural functions of forensic accounting. *Energy Procedia*, 5, 2147–2151. [\[CrossRef\]](#)
- Roszkowska, P. (2021). Fintech in financial reporting and audit for fraud prevention and safeguarding equity investments. *Journal of Accounting & Organizational Change*, 17(2), 164–196. [\[CrossRef\]](#)
- Saifudin, S., Januarti, I., & Purwanto, A. (2025). The role of artificial intelligence in the audit process and how to fraud detections: A literature outlook. *Journal of Ecohumanism*, 4(1), 4185–4203. [\[CrossRef\]](#)
- Sathe, A. K. (2025). Forensic accounting in the age of AI: Detecting fraud with data anomalies. *International Journal of Social Impact*, 10(3), 53–65. [\[CrossRef\]](#)
- Silahtaroglu, G., Dereköy, F., & Baytören, E. (2026). A machine learning approach to audit modification risk prediction in financial reporting: Methods, data, and human-centered challenges. *Journal of Risk and Financial Management*, 19(3), 221. [\[CrossRef\]](#)
- Singleton, T. W., & Singleton, A. J. (2010). *Fraud auditing and forensic accounting* (4th ed.). John Wiley & Sons.
- Sun, H., Zhu, X., & Li, J. (2026). Revealing corporate accounting fraud: From the perspective of individual investors. *Research in International Business and Finance*, 86, 103385. [\[CrossRef\]](#)
- Tavares, F., Pereira, L., Santos, M., & Cunha, P. R. (2023). Challenges of education in the accounting profession in the Era 5.0. *Cogent Business & Management*, 10, 2220198. [\[CrossRef\]](#)
- Tragouda, M., Doumpos, M., & Zopounidis, C. (2024). Identification of fraudulent financial statements through a multi-label classification approach. *Intelligent Systems in Accounting, Finance and Management*, 31(2), e1564. [\[CrossRef\]](#)
- Vitali, S., & Giuliani, M. (2024). The impact of robotic process automation and artificial intelligence on audit firms: Evidence from Italy. *International Journal of Auditing*, 28(1), 45–62. [\[CrossRef\]](#)
- Wolfe, D. T., & Hermanson, D. R. (2004). The fraud diamond: Considering the four elements of fraud. *The CPA Journal*, 74(12), 38–42.
- Zandi, G., Abidin, S. Z., Bakar, L. J. A., & Rehan, R. (2025). Internal financial control enhancement through integration of blockchain and machine learning. *Journal of Information Technology Management*, 17(4), 88–98. [\[CrossRef\]](#)
- Zhou, J., Wang, X., Wang, J., Ye, H., Wang, H., Zhou, Z., Han, D., Ying, H., Wu, J., & Chen, W. (2023). FraudAuditor: A visual analytics approach for collusive fraud in health insurance. *IEEE Transactions on Visualization and Computer Graphics*, 29(6), 2849–2861. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.