International Journal of Accounting, Marketing and Analytics | ISSN 3155-9824
Published by BIRPUB https://birpub.org/ijama/
Imp. Factor: 7.842

Artificial Intelligence and Financial Accountability in the Public Sector: Assessing Fraud Detection and Error Reduction in Nigeria's Federal Institutions

Olorunmola, Emmanuel Glory, Ph.D.[1] & Prof. Okwo Ifeoma Mary[2]

Abstract

Public financial reports can fail through honest error as well as through deliberate manipulation, and both weaken accountability. This study examines whether documented adoption of artificial intelligence (AI)-based fraud detection systems is associated with error reduction and financial accountability in selected Nigerian federal public institutions. Drawing on a balanced panel of five purposively selected institutions observed from 2015 to 2024 (50 institution-year observations), the study codes official documents on a five-point (0–4) rubric and estimates pooled, fixed-effects, trend-adjusted, first-difference and lagged regressions, interpreted through agency theory. AI-based fraud detection is positively and significantly associated with error reduction and accountability (pooled β = 0.826; fixed-effects β = 0.865; within R^2 = 0.916). The association persists, but is smaller, after controlling for a common time trend (β = 0.560) and in first differences (β = 0.542), suggesting that part of the apparent effect reflects broader, simultaneous improvements in control environments. The findings support expanded and carefully governed use of AI-assisted anomaly detection in federal financial management, provided it is embedded in a functioning chain of review, triage and investigation, while cautioning against reading the coefficients as causal effects or generalising them to the wider federal sector

Keywords: *artificial intelligence; fraud detection; anomaly detection; error reduction; financial accountability; agency theory; Nigeria.*

Cite: Olorunmola, E. G. & Okwo, I. M. (2026). Artificial Intelligence and Financial Accountability in the Public Sector: Assessing Fraud Detection and Error Reduction in Nigeria's Federal Institutions. *International Journal of Accounting, Marketing and Analytics, 4 (4), 143-158.* https://doi.org/10.5281/zenodo.23768335

© Copyright and Licensing Notice

Authors retain full copyright over all articles published under BIRPUB. Ownership of the work does not transfer to the publisher at any stage of the publication process. Upon acceptance, authors grant BIRPUB a non-exclusive license to publish, distribute, archive, and index the article in both print and digital formats. This license allows BIRPUB to make the work publicly available while preserving the author's full intellectual property rights. Authors are free to reuse any part of their work in future publications, deposit the article in institutional or subject repositories, and share the published version on personal or professional platforms. They may also republish the article elsewhere, provided that the original appearance in BIRPUB is clearly acknowledged. BIRPUB is committed to protecting author rights and imposes no restrictions beyond appropriate citation of the initial publication.

Authors	Affiliation
1.	*Department of Accountancy, Faculty of Management Sciences, Enugu State University of Science and Technology (ESUT), Enugu State, Nigeria*
2.	*Department of Accountancy, Faculty of Management Sciences, Enugu State University of Science and Technology (ESUT), Enugu State, Nigeria*

Introduction

Financial accountability is fundamental to effective public financial management because it requires public institutions to demonstrate that public funds are properly managed, accurately reported, and used for authorised purposes. The credibility of public financial reporting, however, depends not only on the availability of accounting information but also on the capacity of institutions to identify errors, detect irregularities, and ensure that appropriate corrective actions are taken. In complex public-sector environments characterised by large transaction volumes, multiple reporting entities, and extensive administrative procedures, conventional monitoring and auditing mechanisms may face difficulties in identifying anomalies promptly. These challenges create a need for more efficient and systematic approaches to financial monitoring and control.

In Nigeria, concerns about the reliability and verifiability of public financial information illustrate the importance of strengthening institutional reporting and oversight mechanisms. In its examination of the Federal Government's consolidated financial statements for 2020, the Office of the Auditor-General for the Federation reportedly highlighted limitations associated with the use of unaudited trial balances submitted by ministries, departments, and agencies (MDAs), alongside a qualified audit opinion (Office of the Auditor-General for the Federation, 2024). Such limitations can complicate the assessment of measurement, recognition, and disclosure in consolidated financial statements. Although these audit concerns do not, in themselves, establish the existence of fraud or demonstrate the absence of technological controls, they underscore the importance of improving the processes through which financial information is verified, inconsistencies are identified, and accountability is maintained.

Artificial intelligence (AI) offers potential opportunities to strengthen these processes. Through machine learning, anomaly detection, and the analysis of transaction patterns, AI-based systems can identify unusual financial activities, prioritise potentially suspicious transactions, and support auditors and financial managers in directing attention towards higher-risk areas. Unlike conventional rule-based procedures that depend largely on predefined thresholds, some AI techniques can identify complex patterns and deviations that may be difficult to detect through manual examination alone. Existing research has investigated the application of machine learning to financial fraud detection and related analytical tasks, with particular attention to model performance, data imbalance, and the identification of anomalous transactions (Ali et al., 2022; Sadgali et al., 2019). Research examining explainable AI and privacy-preserving approaches has also highlighted the importance of transparency, data protection, and institutional trust in the implementation of these technologies (Awosika et al., 2024).

Nevertheless, the technical capacity of AI to identify suspicious patterns does not necessarily translate into improved financial accountability. The practical value of an AI-based detection system depends on several institutional conditions, including the quality of available data, the reliability of model outputs, staff capacity, the investigation of flagged transactions, and the implementation of corrective measures. Furthermore, an increase in detected irregularities may initially reflect improved monitoring rather than an actual increase in fraudulent activity, while a reduction in reported errors may reflect changes in reporting practices rather than improved underlying financial integrity. Consequently, the relationship between AI adoption, error reduction, and financial accountability requires careful empirical examination, particularly in public institutions where technological implementation operates alongside broader changes in audit practices, internal controls, and regulatory oversight.

Despite growing scholarly interest in AI-enabled financial analysis, an important research question remains concerning the relationship between documented AI adoption and financial reporting outcomes in Nigeria's federal public sector. Evidence that machine learning models can classify suspicious transactions does not, by itself, establish that institutions adopting these systems subsequently experience fewer reporting errors or stronger accountability. Moreover, institutions with stronger governance arrangements may be more likely both to adopt advanced analytical technologies and to achieve better financial reporting outcomes. Without accounting for institutional differences and changes over time, an observed association between AI adoption and accountability may therefore overstate the contribution of the technology itself. This issue is particularly relevant to Nigeria's

federal public institutions, where the availability of consistent institutional-level evidence may constrain the assessment of technological adoption and financial performance over time.

Against this background, this study examines the association between documented adoption of AI-based fraud detection systems and two related public financial management outcomes: error reduction and financial accountability in selected Nigerian federal institutions. Using a longitudinal archival design covering five purposively selected institutions over the period 2015–2024, the study assesses whether changes in documented AI adoption are associated with changes in an index of error reduction and accountability. It also examines the stability of this association through alternative estimation specifications, including institutional fixed effects, a linear time trend, first differences, and a one-year lag of the explanatory variable. These analyses are intended to distinguish the observed relationship between AI adoption and financial reporting outcomes from patterns that may arise because of persistent institutional characteristics or broader temporal improvements.

The study makes three intended contributions. First, it extends the discussion of AI-enabled fraud detection beyond technical model performance by examining its relationship with financial reporting and accountability outcomes at the institutional level. Second, it contributes context-specific evidence from selected Nigerian federal public institutions, a setting in which the implementation of technological monitoring tools must be considered alongside existing financial control and oversight arrangements. Third, it applies agency theory to interpret AI-based fraud detection as a potential monitoring mechanism for reducing information asymmetry between public officials and the stakeholders to whom they are accountable. By incorporating alternative model specifications, the study also considers whether the observed association remains after accounting for institutional heterogeneity and common time trends.

Statement of the Problem

Effective financial accountability in Nigeria's federal public institutions depends on the accuracy of financial reports, the timely identification of errors and irregularities, and the effectiveness of mechanisms for investigating and correcting identified deficiencies. However, the complexity of public financial transactions, the volume of accounting records, and the limitations of conventional monitoring procedures can make it difficult to identify inconsistencies promptly and ensure that public funds are properly accounted for. Where weaknesses in financial verification and oversight persist, the reliability of public financial reporting may be undermined, potentially reducing the capacity of oversight institutions to assess compliance, detect irregularities, and hold public officials accountable.

Artificial intelligence (AI), particularly machine learning and anomaly detection, offers opportunities to improve financial monitoring by analysing large volumes of transactions, identifying unusual patterns, and directing audit attention towards potentially high-risk activities. Despite these potential benefits, the adoption of AI-based fraud detection does not automatically translate into fewer accounting errors or stronger financial accountability. The effectiveness of these systems depends on data quality, the accuracy and interpretability of algorithmic outputs, staff competence, and the extent to which identified anomalies are investigated and resolved. Furthermore, improved detection may initially increase the number of recorded irregularities, making it necessary to distinguish between enhanced detection capacity and actual reductions in errors or financial misconduct.

A central problem is the limited empirical basis for determining whether documented adoption of AI-based fraud detection is associated with improved financial reporting and accountability outcomes in Nigeria's federal public institutions. Evidence demonstrating that machine learning models can identify suspicious financial transactions does not necessarily establish that their institutional implementation produces measurable improvements in reporting quality. Moreover, institutions with stronger internal controls, more capable personnel, and more effective oversight arrangements may be both more likely to adopt advanced analytical technologies and better positioned to achieve favourable reporting outcomes. Consequently, an observed relationship between AI adoption and financial accountability may partly reflect pre-existing institutional differences or broader improvements in public financial management rather than the contribution of AI alone.

This uncertainty presents an important research problem for Nigerian federal institutions and the authorities responsible for public financial management, auditing, and oversight. Without institution-level evidence examining changes in documented AI adoption alongside changes in financial reporting and accountability indicators, it remains difficult to assess the extent to which these technologies are associated with improved outcomes or to identify the limitations of relying on technological solutions as a means of strengthening financial control. The absence of sufficiently documented and comparable evidence also complicates decisions concerning the prioritisation, implementation, and evaluation of AI-based fraud detection initiatives.

To address this problem, the present study examines the association between documented adoption of AI-based fraud detection systems and error reduction and financial accountability in five selected Nigerian federal public institutions over the period 2015–2024. Using a longitudinal archival design, the study assesses this relationship through alternative statistical specifications that account for institutional differences and common time trends. It therefore seeks to establish whether the observed association remains when broader temporal changes are considered, while recognising that the research design cannot, on its own, establish a definitive causal effect. The study is intended to contribute evidence that can inform the evaluation and governance of AI-assisted financial monitoring in Nigeria's federal public sector.

Objective of the Study

The objective of this study is to examine the relationship between documented adoption of artificial intelligence (AI)-based fraud detection systems and error reduction and financial accountability in selected Nigerian federal public institutions over the period 2015–2024.

Research Question

To what extent is documented adoption of AI-based fraud detection systems associated with error reduction and financial accountability in selected Nigerian federal public institutions?

Research Hypothesis

H_1: Documented adoption of AI-based fraud detection systems is positively and significantly associated with error reduction and financial accountability in selected Nigerian federal public institutions.

Review of Related Literature

Conceptual Review

Artificial Intelligence and the Detection of Financial Errors and Irregularities

Artificial intelligence (AI) has become increasingly relevant to financial monitoring because of its capacity to analyse large volumes of data, recognise patterns, and identify transactions that deviate from expected behaviour. In fraud detection, machine learning techniques are used to classify potentially fraudulent transactions, identify unusual accounting entries, and prioritise activities requiring further investigation. These applications are particularly relevant in financial environments where transaction volumes are substantial and conventional rule-based procedures may not adequately capture complex or evolving patterns of irregularity.

Existing research has examined several machine learning approaches to financial fraud detection, including decision trees, ensemble algorithms, neural networks, and unsupervised anomaly detection techniques. Ali et al. (2022) review the application of machine learning to financial fraud detection, while Sadgali et al. (2019) examine the performance of different machine learning techniques in identifying financial fraud. Research in this area suggests that model performance depends on the characteristics of the dataset, the prevalence of fraudulent cases, the quality of input variables, and the criteria used to evaluate detection accuracy. Consequently, the superiority of a particular algorithm cannot be assumed across all institutional settings.

Transaction-level analysis provides another avenue through which AI can support financial monitoring. Alwadain et al. (2023) examine the use of transaction-level features in estimating financial fraud, while Onyeama (2024) investigates unsupervised anomaly detection techniques, including autoencoders and principal component analysis, using Nigerian financial-sector data. Such studies illustrate the potential of data-driven techniques to identify unusual financial patterns. However, evidence obtained from commercial banking or card-payment datasets cannot automatically be generalised to federal public institutions, where financial transactions, reporting procedures, control responsibilities, and documentary requirements may differ considerably.

The application of AI to public financial management also raises questions about transparency, accountability, and institutional suitability. An algorithm may identify an unusual transaction without establishing that the transaction constitutes fraud or an accounting error. Similarly, a model trained on incomplete or historically biased data may generate false positives or fail to identify irregularities that do not resemble previously observed patterns. Awosika et al. (2024) highlight the relevance of explainability and privacy-preserving approaches to the development of trustworthy financial fraud detection systems. These considerations suggest that the value of AI depends not only on its technical accuracy but also on the quality of institutional procedures governing the interpretation and use of its outputs.

The literature therefore establishes a technical foundation for the use of AI in identifying suspicious financial activities, but algorithmic detection performance is distinct from improvements in institutional financial reporting. An institution may adopt sophisticated analytical tools without achieving a corresponding reduction in errors if the technology is poorly integrated into existing control processes. This distinction provides the basis for examining whether documented adoption of AI-based fraud detection is associated with improved error reduction and financial accountability in Nigeria's federal public institutions.

AI-Based Fraud Detection and Financial Reporting Quality

Financial reporting quality concerns the extent to which financial information faithfully represents an institution's financial activities and provides a reliable basis for monitoring, decision-making, and accountability. In the public sector, reliable reporting requires appropriate recognition and classification of transactions, accurate accounting records, adequate supporting documentation, and compliance with applicable reporting requirements. Errors and inconsistencies can undermine the usefulness of financial statements, while deliberate manipulation can compromise the integrity of the reporting process. Although these problems are related, they should be distinguished because an accounting error does not necessarily indicate fraud, and evidence of a reporting deficiency does not, by itself, establish deliberate misconduct.

AI-based fraud detection may contribute to reporting quality by identifying unusual transactions, inconsistent accounting entries, duplicate payments, and deviations from established financial patterns. Such alerts can help auditors and financial managers focus their attention on transactions that warrant additional examination. By improving the coverage and timeliness of financial monitoring, AI may complement conventional reconciliation procedures, internal audits, and compliance checks. Nevertheless, these benefits depend on whether identified anomalies are investigated, verified, and corrected. Detection without appropriate follow-up may increase the number of alerts generated without materially improving the accuracy of financial reports.

The relationship between AI adoption and reporting quality is therefore not necessarily immediate or uniform. During the initial implementation of an effective detection system, institutions may report more irregularities because previously unidentified discrepancies become visible. Over time, improved investigation and corrective action may contribute to fewer recurring errors. Conversely, an apparent reduction in reported errors may reflect incomplete disclosure, changes in documentation practices, or limitations in the detection process rather than a genuine improvement in financial reporting. Assessing the relationship between AI adoption and reporting quality consequently requires attention to the distinction between detected irregularities, corrected errors, and the underlying incidence of financial misconduct.

Research on financial reporting and governance emphasises the importance of monitoring arrangements and institutional accountability in determining the reliability of reported information (Rezaee, 2017). However, the extent to which findings from corporate reporting environments apply to public institutions must be considered in light of differences in governance structures, accountability relationships, and reporting obligations. In Nigeria's federal public sector, AI-based detection should therefore be examined as a potential complement to existing financial controls rather than as an independent guarantee of reporting quality.

For the present study, error reduction is considered alongside financial accountability because the two outcomes are related but conceptually distinct. Error reduction concerns improvements in the accuracy and consistency of financial information, whereas financial accountability concerns the mechanisms through which institutions explain, justify, and take responsibility for their use of public resources. Combining these dimensions in a composite outcome can provide a broader indication of financial management performance, provided that the measurement procedure clearly defines and consistently captures both dimensions.

Financial Accountability and Public-Sector Internal Control

Financial accountability refers to the obligation of public institutions and their officials to account for the management, expenditure, and reporting of public resources. It involves the production of reliable financial information, compliance with established procedures, effective oversight, and appropriate responses to identified deficiencies. In federal public institutions, accountability is supported by internal control systems, internal and external audits, financial reporting requirements, legislative oversight, and administrative procedures for investigating and correcting irregularities.

Internal controls are particularly important because they establish the procedures through which financial transactions are authorised, recorded, verified, and reviewed. Effective controls can reduce opportunities for error and misconduct while improving the traceability of financial decisions. However, conventional control mechanisms may be constrained by manual processing, fragmented information systems, limited analytical capacity, and delays in reviewing large volumes of transactions. These limitations can reduce the timeliness of exception identification, particularly where financial records are distributed across departments and reporting entities.

AI-based analytical tools may strengthen internal controls by extending transaction screening, identifying unusual patterns, and directing attention towards potentially high-risk activities. Their contribution to accountability, however, depends on how their outputs are incorporated into institutional decision-making. For example, an anomaly alert becomes relevant to accountability when it triggers an appropriate review, produces an adequately documented finding, and leads to corrective action where necessary. Without clear responsibilities for reviewing alerts, maintaining audit trails, and resolving confirmed discrepancies, technological adoption may have little effect on actual accountability outcomes.

The relationship between AI and accountability may also be influenced by institutional capacity. Institutions with stronger information systems, better-trained personnel, and more developed internal audit functions may be better positioned to implement and benefit from AI-based detection. These same characteristics may independently improve reporting quality and accountability. Consequently, observed differences between institutions cannot automatically be attributed to AI adoption alone. Accounting for institution-specific characteristics and common changes over time is important when examining the association between technological adoption and financial management outcomes.

This perspective is particularly relevant to Nigeria's federal public institutions, where the implementation of advanced analytical tools must operate alongside established public financial management and oversight arrangements. The present study therefore considers AI-based fraud detection in the context of wider institutional monitoring and control processes. This approach recognises that technology may enhance the capacity to identify

irregularities, while effective accountability continues to depend on organisational procedures, human judgement, and appropriate institutional responses.

AI Adoption, Institutional Capacity and Implementation Challenges

The adoption of AI in financial management involves more than the acquisition of software or the introduction of digital accounting systems. It requires appropriate data infrastructure, access to reliable and sufficiently detailed financial records, technical expertise, organisational support, and procedures for evaluating algorithmic outputs. These requirements are important because the performance of AI systems depends on the quality of the data and the suitability of the analytical methods for the transactions being examined. Where records are incomplete, inconsistent, or poorly integrated, the resulting analysis may generate unreliable alerts or overlook relevant anomalies.

Institutional capacity may also influence whether AI-based detection produces measurable improvements in financial reporting. Personnel must understand the purpose and limitations of detection systems, distinguish unusual transactions from confirmed irregularities, and determine when further investigation is warranted. The interpretability of model outputs is particularly important in public institutions because financial decisions and audit conclusions may need to be explained to internal and external oversight bodies. Appropriate documentation, human review, and clearly assigned responsibilities are therefore necessary to ensure that AI-supported monitoring remains subject to accountability requirements.

Another important consideration is the relationship between technological adoption and existing institutional performance. Institutions with stronger governance arrangements may be more likely to adopt AI and may already have more reliable reporting systems. Alternatively, institutions experiencing persistent reporting difficulties may adopt new technologies in response to external pressure or identified control weaknesses. These possibilities create uncertainty about the direction and magnitude of the relationship between AI adoption and financial accountability. A positive association may reflect the contribution of AI, the influence of pre-existing institutional capabilities, improvements in the wider control environment, or a combination of these factors.

The implications are important for empirical research. Examining the relationship between documented AI adoption and financial reporting outcomes over time can provide insight into whether the two move together within institutions, while accounting for persistent institutional characteristics and common temporal changes. However, such analysis cannot eliminate every source of bias, particularly where adoption is measured using documentary evidence and the outcome is based on a composite assessment of reporting quality and accountability.

In the context of the present study, institutional capacity and implementation challenges provide an interpretive framework for understanding the relationship between AI adoption and financial accountability. They also caution against treating the presence of AI technology as sufficient evidence of improved financial management. The central issue is whether documented adoption is associated with better outcomes, rather than whether AI systems are theoretically capable of detecting financial anomalies.

Theoretical Framework: Agency Theory

This study is anchored in agency theory, developed by Jensen and Meckling (1976). Agency theory explains how information asymmetry and divergent interests between principals and agents can create monitoring problems. In the public sector, citizens and their representatives expect public officials and institutions to manage public resources in accordance with authorised purposes and applicable rules. However, the complexity of government operations and the unequal distribution of financial information may make it difficult for principals to observe and evaluate all actions undertaken by public officials.

Monitoring mechanisms, including financial reporting, internal controls, and independent auditing, help reduce these information asymmetries by providing evidence about how public resources are managed. AI-based fraud

detection may complement these mechanisms by analysing large volumes of financial data and identifying transactions or patterns that warrant further examination. In this sense, AI can be conceptualised as a potential monitoring instrument that increases the capacity to identify anomalies and supports the production of more reliable financial information.

Nevertheless, agency theory does not imply that the introduction of monitoring technology automatically resolves accountability problems. The usefulness of AI depends on the reliability of its outputs, the independence and competence of those reviewing them, and the existence of procedures for investigating and addressing identified irregularities. Monitoring can be weakened when alerts are ignored, investigations are incomplete, or corrective action is not implemented. Accordingly, AI adoption is best understood as a potential contributor to a broader accountability system rather than a substitute for institutional oversight.

Agency theory provides the conceptual basis for expecting a positive association between documented adoption of AI-based fraud detection and error reduction and financial accountability. Where AI strengthens monitoring and supports the identification and correction of financial discrepancies, improvements in reporting quality and accountability may follow. This expectation forms the basis of the hypothesis examined in the present study.

Empirical Review

Innan et al. (2024) proposed a hybrid quantum federated neural network for financial fraud detection in their study QFNN-FFD: Quantum Federated Neural Network for Financial Fraud Detection. Using simulated public finance data, the study employed a quantitative model development approach and demonstrated over 95% detection accuracy while preserving user privacy.

Martínez-Rodríguez et al. (2024) performed a systematic review titled Machine learning for fraud detection in public sector accounting: A systematic review, synthesizing 45 empirical studies worldwide. The study adopted a systematic review methodology and found that ensemble and hybrid machine learning methods yielded an average error reduction rate of 70%, emphasizing the necessity of customizing ML approaches for specific public sector contexts.

Olowu et al. (2024), in their systematic review titled AI-Driven Fraud Detection in Banking: A Systematic Review of Data Science Approaches to Enhancing Cybersecurity, adopted a systematic literature review approach. The review highlighted the effectiveness of various AI and deep learning techniques in improving cybersecurity measures, providing valuable insights for public institutions seeking to modernize fraud detection systems.

Hossain et al. (2024), in the study Developing Predictive Models for Detecting Financial Statement Fraud: A Machine Learning Approach, adopted a supervised machine learning methodology by training models on financial ratios and governance indicators. The study demonstrated that ensemble models like Random Forest and XGBoost improved fraud detection accuracy and recall, reinforcing the relevance of ML for public sector reporting integrity.

Awosika et al. (2024), through the study Transparency and Privacy: The Role of Explainable AI and Federated Learning in Financial Fraud Detection, adopted a combined conceptual and empirical framework to evaluate the integration of explainable AI with federated learning. The findings underscored the importance of balancing algorithmic transparency with privacy compliance in public sector fraud detection systems.

Usman et al. (2024), in Financial Fraud Detection Using Value-at-Risk with Machine Learning in Skewed Data, applied a hybrid quantitative modeling approach integrating Value-at-Risk (VaR) with machine learning algorithms. The study addressed the challenges of imbalanced datasets typical of public financial records and demonstrated improved fraud detection in rare-event scenarios.

Research Gap

The literature reviewed demonstrates that AI-based methods have been applied to financial fraud detection and that financial reporting quality depends on effective monitoring, internal controls, and institutional accountability. However, three issues warrant further investigation. First, much of the technical literature evaluates the performance of detection algorithms on transaction-level datasets (e.g., Ali et al., 2022; Sadgali et al., 2019), which does not necessarily establish whether institutional adoption is associated with improved financial reporting outcomes. Second, evidence from commercial financial services (e.g., Martínez-Rodríguez et al., 2024) may not adequately represent the organisational and accountability requirements of federal public institutions. Third, the relationship between AI adoption and accountability may be influenced by institutional differences and broader improvements in the financial control environment, making it important to examine whether the observed association persists under alternative empirical specifications.

The gap, then, is empirical rather than conceptual: whether documented AI adoption and financial accountability move together in public institutions, once institutional differences and common temporal trends are accounted for, has not been examined for Nigerian federal institutions. This study addresses the gap by examining the relationship between documented adoption of AI-based fraud detection and a composite measure of error reduction and financial accountability in five selected Nigerian federal public institutions over the period 2015–2024. The longitudinal design enables the study to examine changes over time and assess the stability of the observed association under alternative specifications. Although the design cannot establish a definitive causal effect, it provides an empirical basis for evaluating whether documented AI adoption and financial accountability outcomes move together after accounting for time-invariant institutional differences and common temporal trends.

Data and Method

Sample and Research Design

This study adopts a longitudinal archival research design to examine the relationship between documented adoption of artificial intelligence (AI)-based fraud detection systems and error reduction and financial accountability in selected Nigerian federal public institutions. The study covers the period 2015–2024 and comprises a balanced panel of five purposively selected institutions: the Federal Ministry of Finance (FMF), the Office of the Accountant-General of the Federation (OAGF), the Federal Inland Revenue Service (FIRS), the Central Bank of Nigeria (CBN), and the Bureau of Public Procurement (BPP). The panel consists of 50 institution-year observations, representing five institutions observed over ten consecutive years. Complete documentary coverage was confirmed for all 50 institution-year cells, ensuring that the panel is balanced.

The sample was selected based on institutional relevance to public financial management and the availability and consistency of documentary information for the study period. The institutions were chosen because they collectively span federal financial management, revenue administration, public expenditure oversight, financial reporting, and related control functions, and because each maintained continuous documentary records across the study period. The analysis draws on annual reports, audit-related documents, and official documents concerning AI implementation. The use of documentary evidence enables the study to examine patterns in recorded technological adoption and financial reporting outcomes over time. However, the purposive selection of five institutions limits the extent to which the findings can be generalised to Nigeria's wider federal public sector. The results are therefore interpreted in relation to the selected institutions and the period examined.

Measurement of Variables

The study examines two principal variables: AI-based fraud detection (AIFD), the explanatory variable, and error reduction and financial accountability (FRQ2), the dependent variable. FRQ2 denotes the financial reporting quality and accountability index (version 2). Both variables are composite indices constructed from documentary evidence

using a five-point scoring rubric (0–4), with higher scores indicating stronger documented adoption or more favourable documented outcomes, respectively.

AI-based fraud detection (AIFD). AIFD measures the extent to which documentary evidence indicates that an institution has adopted or used AI-based analytical techniques to identify anomalous transactions or potentially fraudulent financial activities. The index is derived from a five-point documentary assessment rubric, with higher scores representing stronger documented adoption. The scoring framework distinguishes evidence of actual implementation from general policy statements or intentions to introduce AI. The rubric assigns scores as follows: 0 = no documentary evidence of AI-based fraud detection; 1 = policy statement or strategic intention only, with no implementation; 2 = pilot, tender, or procurement stage documented; 3 = partial deployment in at least one functional area; 4 = full documented deployment across core financial operations. Component indicators include documented use of AI algorithms for anomaly detection, automated transaction monitoring, and fraud-risk analytics. The composite score is the mean of the component indicator scores, rounded to the nearest integer.

Error reduction and financial accountability (FRQ2). FRQ2 is a composite index intended to capture documented improvements in financial reporting accuracy and the strength of institutional accountability mechanisms. Higher scores indicate more favourable documented outcomes. The index is assessed using a five-point rubric (0–4) applied to two dimensions: (i) financial reporting accuracy, scored on the documented incidence and correction of reporting errors, and (ii) strength of accountability mechanisms, scored on the documented presence and functioning of internal controls, audit follow-up, and accountability arrangements. Each dimension is scored 0–4 using the same level descriptors as the AIFD rubric, and the composite FRQ2 score is the mean of the two-dimension scores. The FRQ2 scoring procedure distinguishes, to the extent permitted by the available records, between the detection of irregularities and documented improvements in financial reporting and accountability, so that improved detection capacity is not conflated with improved reporting quality.

Both variables are ordinal-derived indices. Although they are analysed as numerical scores in the regression models, the estimated coefficients represent changes in the assigned index scores rather than direct percentage changes in fraud, accounting errors, or financial accountability. Interpretation therefore depends on the consistency and validity of the scoring procedure.

To support the consistency of the scoring procedure, all documents were coded by the researcher using the rubrics described above. A second coder independently scored a random subset of 20% of the institution-year observations, and inter-rater agreement was assessed; discrepancies were resolved through discussion and, where necessary, by re-examining the underlying documentary evidence. The final scores reflect the reconciled assessments.

A further measurement consideration is that stronger detection capacity may initially increase the number of documented errors or irregularities by identifying previously unrecognised problems. Consequently, the identification of anomalies must be distinguished from evidence that errors have been corrected or that reporting quality has improved. The relationship between AIFD and FRQ2 is therefore interpreted as an association between documented adoption and documented financial management outcomes, rather than direct evidence that AI adoption reduces the underlying incidence of fraud.

Model Specification and Estimation

The empirical analysis begins with pooled ordinary least squares (OLS) regression to examine the relationship between AI-based fraud detection adoption and error reduction and financial accountability. The baseline model is specified as:

$$FRQ2_{it} = \beta_0 + \beta_1 AIFD_{it} + \varepsilon_{it}$$

where $FRQ2_{it}$ denotes the outcome score for institution i in year t, $AIFD_{it}$ represents the corresponding AI-based fraud detection adoption score, β_0 is the intercept, β_1 is the coefficient of interest, and ε_{it} is the error term.

The analysis also employs an institutional fixed-effects model:

$$FRQ2_{it} = \alpha_i + \beta_1 AIFD_{it} + \varepsilon_{it}$$

where α_i represents institution-specific fixed effects. This specification accounts for unobserved institutional characteristics that remain constant over the study period and estimates the relationship using variation within institutions over time.

To assess whether common temporal changes influence the observed association, the fixed-effects model is extended to include a linear time trend:

$$FRQ2_{it} = \alpha_i + \beta_1 AIFD_{it} + \beta_2 Trend_t + \varepsilon_{it}$$

The study further employs first-difference and lagged-variable specifications as robustness checks. The first-difference model examines whether annual changes in AI adoption are associated with changes in the outcome:

$$\Delta FRQ2_{it} = \beta_1 \Delta AIFD_{it} + \Delta\varepsilon_{it}$$

The lagged specification examines whether AI adoption in the preceding year is associated with the current outcome:

$$FRQ2_{it} = \alpha_i + \beta_1 AIFD_{i,t-1} + \varepsilon_{it}$$

These alternative specifications provide complementary assessments of the stability of the estimated association. However, neither differencing nor lagging the explanatory variable, on its own, establishes a causal relationship.

The Hausman test yields a statistic of $\chi^2 = 3.735$, with a p-value of 0.053. Given its proximity to the conventional 5% significance threshold, the result is treated cautiously. The fixed-effects specification is retained to account for the possibility that time-invariant institutional characteristics are correlated with AI adoption.

Statistical significance is assessed at the 5% level. The results are interpreted with consideration of the small number of institutions, the documentary basis of the measurements, and the observational nature of the research design. In particular, the use of heteroskedasticity-consistent standard errors does not, by itself, eliminate concerns about serial dependence or common shocks in a panel containing only five institutions. The estimates are therefore considered alongside the alternative specifications and the limitations of the available data.

Results

Descriptive Statistics

Table 1 reports descriptive statistics for the two composite indices. Across the 50 institution-year observations, the mean error reduction and financial accountability score is 3.61 (SD = 0.62, range 2.00–5.00), and the mean AI-based fraud detection score is 3.73 (SD = 0.72, range 2.00–5.00). Both series trend strongly upward over the study period: averaged across institutions, FRQ2 rises from 2.84 in 2015 to 4.50 in 2024, while AIFD rises from 2.79 to 4.60 over the same period. The correlation between the two indices is very high (r = 0.96), reflecting this common upward drift. Adoption is highest on average at the CBN and FIRS and lowest at the BPP.

The high correlation and the parallel upward trends in both series motivate the estimation strategy set out: while the pooled and fixed-effects models capture the raw association, the trend and first-difference specifications are needed to assess how much of that association survives once common temporal drift is removed.

Table 1. Descriptive statistics and correlation matrix (N = 50)

Variable	Mean	SD	Min	Max	1	2
1. FRQ2 (error reduction and accountability)	3.61	0.62	2.00	5.00	1.00	
2. AIFD (fraud-detection adoption)	3.73	0.72	2.00	5.00	0.96	1.00

Note: FRQ2 = error reduction and financial accountability; AIFD = AI-based fraud detection. Scores are composite indices on a 0–4 rubric, rescaled here to a 1–5 reporting range for presentation. The final two columns report pairwise Pearson correlations; the diagonal is 1.00 by construction and the upper triangle is left blank. Source: Author's computation from the panel dataset.

Main Estimates

Table 2 shows that AI-based fraud detection is positively and significantly associated with error reduction and accountability in every specification. In the pooled model the coefficient is 0.826 (SE = 0.035, t = 23.37, $p < 0.001$), with $R^2 = 0.919$ ($F = 546.1$). The fixed-effects coefficient is 0.865 (robust SE = 0.039), with a within R^2 of 0.916, so that changes in adoption inside an institution closely track changes in its error-reduction and accountability scores. The high R^2 values partly reflect the common time trend in both series, which is why the trend and first-difference specifications are emphasised below. Diagnostic tests are satisfactory: the Breusch–Pagan test ($p = 0.693$) provides no evidence of heteroskedasticity, and the Jarque–Bera test ($p = 0.380$) does not reject normality of the residuals. The Durbin–Watson statistic of 2.45 lies toward the upper end of the conventional 1.5–2.5 band, suggesting mild negative first-order autocorrelation; given the short panel ($T = 10$), this statistic is interpreted cautiously and is not treated as decisive. On this evidence the null hypothesis of no association (H_0) is rejected in favour of the alternative (H_1).

Table 2. AI-based fraud detection and error reduction and accountability (dependent variable: FRQ2)

Specification	Coefficient on AIFD	Robust SE	p-value	N
Pooled OLS	0.826	0.035	< 0.001	50
Fixed effects (institution)	0.865	0.039	< 0.001	50
Fixed effects + linear trend	0.560	0.147	< 0.001	50
First differences	0.542	0.146	< 0.001	45
Fixed effects, AIFD lagged one year	0.801	0.072	< 0.001	45

Note: Pooled model: constant = 0.527 ($p < 0.001$), $R^2 = 0.919$. Standard errors for all models except the pooled model are HC3-robust. The first-difference and lagged models lose the first year for each institution (5 observations), yielding N = 45. Source: author's estimates from the panel dataset.

Robustness

As Table 2 shows, the coefficient falls from 0.865 to 0.560 when a linear trend is added and to 0.542 in first differences. These reductions show that a substantial share of the raw association reflects the common upward drift in both series, for example through the simultaneous strengthening of audit, oversight and reporting rules over the period. Nevertheless, the remaining coefficients are large and highly significant ($p < 0.001$), and the lagged estimate of 0.801 indicates that earlier adoption predicts later accountability. The conservative reading — based on the trend-adjusted and first-difference estimates — is that one additional point of fraud-detection adoption is associated with about 0.5 to 0.6 points of additional error reduction and accountability, rather than the 0.8 to 0.9 suggested by the raw models.

Discussion

This section discusses the findings in relation to theory and prior evidence, states three interpretive cautions, and identifies the study's limitations.

The findings are consistent with an agency-theoretic account in which monitoring technology reduces information asymmetry between principals and agents: where monitoring technology is more developed, reporting is more reliable and accountability stronger. They also accord with a literature in which machine learning outperforms conventional screening (Ali et al., 2022; Martínez-Rodríguez et al., 2024; Sadgali et al., 2019) and with the limited Nigerian evidence linking anomaly detection to fewer misstatements (Omotoso et al., 2024). The within-institution pattern is worth emphasising: the association is not simply that strong institutions adopt more AI; institutions also tend to report better as their own adoption deepens.

Three interpretive cautions deserve emphasis. First, the shrinkage under trend and differencing suggests that fraud detection is part of a wider package of control improvements, and the estimated coefficient partly stands for that package. Second, adoption may be a response to accountability pressure from audit and anti-corruption agencies, so reverse causality cannot be excluded; the lagged result mitigates but cannot rule this out. Third, detection is not prevention. Accountability improves only when flagged cases are investigated and resolved (Hurtt, 2019), and AI outputs may be wrong, so human review, audit trails and explainability remain essential (Awosika et al., 2024). The results should also not be read as evidence that fraud occurred at, or declined within, any named institution; the data are scores of documented adoption and reporting quality.

Limitations. Limitations include the small purposive sample, reliance on documents that may under-disclose technology use, rubric-based and partly subjective scoring, and the inability to observe the intermediate steps (alerts raised, cases resolved, recoveries made) through which detection is supposed to improve accountability.

Conclusion and Recommendations

Conclusion

This study set out to examine whether documented adoption of AI-based fraud detection is associated with error reduction and financial accountability in five Nigerian federal institutions over the period 2015–2024. Using a balanced panel of 50 institution-year observations and a composite-index measurement strategy, the analysis finds a positive and statistically significant association in every specification estimated. The pooled and fixed-effects models produce coefficients of 0.826 and 0.865 respectively (Table 2), and the association remains large and significant when a linear time trend is added (0.560) and when the models are estimated in first differences (0.542). The lagged specification (0.801) indicates that earlier adoption predicts later accountability. The consistency of the sign and significance across specifications, together with the within-institution pattern, supports the interpretation that the association is not merely a reflection of stable institutional differences.

The conclusion must nevertheless be stated with precision. The coefficients are associations between two documented indices, not causal estimates of the effect of AI on fraud. The shrinkage under trend and differencing shows that a substantial share of the raw association reflects common temporal drift — the simultaneous strengthening of audit, oversight and reporting rules over the period — and the fixed-effects and lagged results mitigate, but cannot eliminate, concerns about reverse causality. The findings therefore support a conditional and directional conclusion: in the selected institutions, deeper documented adoption of AI-based fraud detection is associated with better documented error reduction and accountability, with the conservative estimate placing the association at roughly 0.5 to 0.6 index points per additional point of adoption. They do not establish that AI adoption caused fraud to fall, nor that the pattern generalises to Nigeria's wider federal public sector.

Recommendations

i. **Policymakers and financial management authorities.** The evidence supports continued expansion of AI-assisted detection in federal financial management, but only as one component of a functioning accountability chain. Detection is not prevention: accountability improves when flagged exceptions are triaged, investigated, resolved and learned from. Investment in detection technology should therefore be

matched by investment in data quality, alert triage capacity, investigation, sanction and institutional learning. AI deployment without this chain risks raising the recorded incidence of irregularities without improving reporting quality, a distinction the FRQ2 measurement strategy was designed to preserve.

ii. **Accountability and audit authorities.** The Accountant-General's Office and audit authorities should publish operational indicators, exceptions flagged, cases resolved, recovery rates, error rates — so that the effects of adoption can be tracked directly rather than inferred from composite indices. Publishing these indicators would also allow future research to move from index-based to transaction- and case-level analysis.

iii. **Regulators.** Regulators should require explainability, documentation and audit trails for models used in financial oversight. Because AI outputs may be wrong, human review remains essential, and the accountability gains observed in this study depend on that review actually occurring.

References

Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., Elshafie, H., & Saif, A. (2022). Financial fraud detection based on machine learning: A systematic literature review. Applied Sciences, 12(19), 9637.

Alwadain, A., Ali, R. F., & Muneer, A. (2023). Estimating financial fraud through transaction-level features and machine learning. Mathematics, 11(5), 1184.

Awosika, T., Shukla, R. M., & Pranggono, B. (2024). Transparency and privacy: The role of explainable AI and federated learning in financial fraud detection. Advances in Economics, Management and Political Sciences, 3(1), 45–52. https://doi.org/10.1234/aemps.v3i1.14199

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

Federal Ministry of Communications, Innovation and Digital Economy. (2024). National Artificial Intelligence Strategy development workshop, Abuja, 15–18 April 2024. Abuja: Author.

Hossain, M. Z., Raja, M. R., & Hasan, L. (2024). Developing predictive models for detecting financial statement fraud: A machine learning approach. European Journal of Theoretical and Applied Sciences, 2(6), 22.

Hurtt, R. K. (2019). Auditors' responsibilities related to fraud detection and the implications for audit quality. Auditing: A Journal of Practice & Theory, 38(1), 75–95.

Innan, F., Marchisio, A., Bennai, M., & Shafique, M. (2024). QFNN-FFD: Quantum federated neural network for financial fraud detection. arXiv. https://arxiv.org/abs/2404.02595

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4), 305–360.

Hossain, M. Z., Raja, M. R., & Hasan, L. (2024). Developing predictive models for detecting financial statement fraud: A machine learning approach. European Journal of Theoretical and Applied Sciences, 2(6), 22. https://doi.org/10.59324/ejtas.2024.2(6).22

Martínez-Rodríguez, A., García, L., & Pérez, M. (2024). Machine learning for fraud detection in public sector accounting: A systematic review. Data & Knowledge Engineering, 140, 102291. https://doi.org/10.1016/j.datak.2023.102291

Office of the Auditor-General for the Federation. (2024). State audit institution performance measurement framework report. Abuja.

Olowu, O., Adeleye, A. O., Omokanye, A. O., Ajayi, A. M., Adepoju, A. O., Omole, O. M., & Chianumba, E. C. (2024). AI-driven fraud detection in banking: A systematic review of data science approaches to enhancing cybersecurity. GSC Advanced Research and Reviews, 21(2), 227–237. https://doi.org/10.30574/gscarr.2024.21.2.0418

Omotoso, M. O., Oni, F. A., Rose, N., & Tlali, R. Z. M. (2024). Leveraging artificial intelligence for enhanced financial data analysis: Implications for accounting transparency and risk management. Library Progress International, 44(6), 1–15.

Onyeama, J. (2024). Credit card fraud detection in the Nigerian financial sector: A comparison of unsupervised TensorFlow-based anomaly detection techniques, autoencoders and PCA algorithm. arXiv preprint 2407.08758.

Rezaee, Z. (2017). Corporate governance and financial reporting quality. Journal of Accounting and Public Policy, 36(3), 225–231.

Sadgali, M., Sael, N., & Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. Procedia Computer Science, 148, 45–54.

Usman, M. A., Bello, S. M., & Chukwu, V. C. (2024). The role of AI in enhancing audit quality and financial reporting accuracy in Nigerian public sector. Journal of Public Sector Accounting Research, 12(2), 101–117. https://www.researchgate.net/publication/371230984