Journal of Information Technology Management
ISSN-OnLine: 2980-7972
University of Tehran

Explainable Machine Learning for Audit Risk Analytics: Digital Disclosure Complexity and Auditor–Client Network Structure

Nazar Maan Abdul Kareem*

*Corresponding author, Department of Accounting, College of Administration and Economics, University of Diyala, 32001, Ba'aqubah, Diyala, Iraq. E-mail: dr.nazarm@uodiyala.edu.iq

Ammar Ghazi Ibrahim Al-Ezzi

Department of Accounting, College of Administration and Economics, University of Diyala, 32001, Ba'aqubah, Diyala, Iraq. E-mail: ammar.ezzi82@uodiyala.edu.iq

Waleed Khamees Sabea

Department of Accounting, College of Administration and Economics, University of Diyala, 32001, Ba'aqubah, Diyala, Iraq. E-mail: waleedkhamees@uodiyala.edu.iq

Journal of Information Technology Management, 2026, Vol. 18, Special Issue, pp. 102-121.
Published by the University of Tehran, College of Management
doi: https://doi.org/10.22059/jitm.2026.109975
Article Type: Research Paper
© Authors
Received: August 13, 2026
Received in revised form: October 24, 2026
Accepted: November 08, 2026
Published online: December 01, 2026

Abstract

This study develops a reproducible, explainable machine-learning framework for examining whether the digital complexity of corporate disclosures and the structure of auditor–client networks are associated with audit-deficiency risk. The framework addresses a gap in information systems research on regulatory analytics by integrating heterogeneous digital filing data into a leakage-safe, interpretable analytical pipeline. The proposed architecture integrates three public data sources: (1) SEC EDGAR/XBRL filings—that is, corporate filings in the U.S. Securities and Exchange Commission's (SEC) Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system, tagged in eXtensible Business Reporting Language (XBRL), a machine-readable format for financial data; (2) PCAOB Form AP auditor–client disclosures; and (3) PCAOB Part I.A inspection outcomes. Descriptive analysis indicates that audits with multiple deficiencies dominate the disclosed data and that revenue, inventory, and receivables are the most frequently coded audit areas. These patterns suggest that audit-deficiency risk is concentrated in specific financial-statement domains and is associated with identifiable structural and disclosure characteristics. The framework specifies leakage-safe temporal validation, calibrated classification, and SHAP (SHapley Additive exPlanations)-based explanations for the completed SEC–Form AP–PCAOB panel. The study contributes a defensible data architecture that distinguishes regulatory inspection evidence from broader

audit-quality claims. It offers a reproducible, explainable AI pipeline for regulatory analytics, with implications for audit oversight, information systems governance, and the design of decision-support tools in regulatory technology.

Keywords: Digital Disclosure Complexity; Audit Quality; PCAOB Inspections; Audit Data Analytics; Explainable Machine Learning

Introduction

Audit quality is fundamentally concerned with whether an audit engagement produces sufficient and appropriate evidence, supports an independent opinion, and contributes to credible financial reporting. Although the construct is multidimensional and cannot be observed perfectly, inspection outcomes, restatements, and financial reporting failures remain crucial empirical signals (Francis, 2011; DeFond & Zhang, 2014). Digital reporting changes the information environment in which this assurance is produced. U.S. registrants now distribute large volumes of structured XBRL facts and narrative disclosures through EDGAR, while audit firms increasingly apply data analytics, artificial intelligence, and technology-assisted procedures. These developments may improve coverage and risk assessment, yet they also make the quality of the underlying data, transformations, and analytical judgment more consequential (Brown-Liburd et al., 2015; Vasarhelyi et al., 2015; Fedyk et al., 2022).\

A central unresolved issue is whether the complexity of digital disclosure itself is associated with audit-deficiency risk. Complexity is not merely the length of a filing. It includes the intensity and diversity of XBRL tagging, reliance on entity-specific taxonomy extensions, narrative density, difficult-to-read language, risk-related disclosures, and potential inconsistencies between numbers and accompanying explanations. These features may reflect economically complex transactions and estimation uncertainty, but they may also introduce information-processing costs for preparers, auditors, regulators, and investors. Audit data analytics could aid auditors in examining larger populations and identifying unusual patterns; however, analytics does not remove the need to establish data reliability, relevance, and a defensible audit trail (Appelbaum et al., 2017; Huang et al., 2022; PCAOB, 2025d).

The auditor's task is further shaped by the structure of its client portfolio. Form AP disclosures make it possible to represent public-company auditing as a bipartite network in which audit firms and engagement partners are connected to issuer clients. Network characteristics such as client concentration, portfolio diversity, partner workload, repeat-client tenure, and component-auditor participation may influence the resources, knowledge, coordination demands, and quality-control challenges facing an audit firm. Existing studies have investigated the informativeness of inspection outcomes and the implications of digital expertise; however, research has rarely integrated machine-readable reporting complexity

with auditor–client network structure and a public regulatory outcome in one explainable predictive framework (Aobdia, 2018, 2019; Rahman & Ziru, 2023).

The research problem is consequently twofold. First, conventional audit-quality research often relies on archival proxies that are distant from the regulatory sufficiency-of-evidence construct. Second, various machine-learning studies predict accounting fraud or restatement events without separating the timing of public disclosure, auditor assignment, and subsequent inspection results. Such designs may generate information leakage or explanations that cannot be translated into audit planning. Recent professional guidance reinforces the need for careful documentation and governance when artificial intelligence or technology-assisted analysis is used in audit work (Financial Reporting Council [FRC], 2025; PCAOB, 2024, 2025c).

Accordingly, this study has three objectives. It first constructs an issuer-level Digital Disclosure Complexity Index (DDCI) from SEC XBRL and filing text. It next builds auditor–client network measures from Form AP, aggregated to the audit-firm-year level, as public PCAOB Part I.A reports do not disclose the identity of each inspected issuer. Finally, it estimates and compares an interpretable logistic benchmark, elastic-net regularization, random forest, and gradient-boosting models for predicting high firm-year Part I.A deficiency risk. Model explanations are generated with SHAP values and reported alongside calibration diagnostics, thereby preventing predictive accuracy from being treated as sufficient evidence of audit relevance.

The empirical results reported in this manuscript are deliberately limited to a data audit of the public PCAOB Part I.A file. The source identifies individual inspection findings and an anonymized issuer reference key, but it does not provide issuer CIKs; consequently, an issuer-level SEC link cannot be inferred reliably from that file alone. This restriction is methodological rather than cosmetic. It prevents an invalid linkage and permits a transparent distinction between observed regulatory results and the completed predictive analysis. The PCAOB itself cautions that inspected audits are not a representative sample of a firm's entire issuer population and that inspection reports are not overall firm ratings (PCAOB, 2026b).

The new contribution is therefore a data-governed, audit-specific framework rather than an unsupported claim that machine learning automatically improves audit quality. The framework combines the digital properties of the client's public reporting, the audit firm's observable network structure, and a regulatory inspection outcome. It imposes a strict temporal order and requires performance, calibration, and explanation to agree before any factor is interpreted as a risk signal. This architecture is intended to support future audit-quality research using public data while preserving the distinction between association, prediction, and causal inference.

Literature Review

Digital disclosure complexity and audit quality

The literature on audit data analytics establishes that structured and unstructured data could strengthen risk assessment, anomaly detection, and population-level testing, but it also emphasizes behavioral, governance, and data-quality constraints. XBRL makes numerous financial statement facts comparable across entities and periods, while issuer-specific extensions and diverse narrative disclosures capture company-specific reporting choices. These characteristics may reveal a complex information environment that requires more audit effort and more sophisticated audit evidence. Rather than treating all longer or more technical reports as risky, the proposed DDCI decomposes complexity into observable dimensions. This approach aligns with prior work on the use of data analytics in auditing and with research showing that client digitalization and audit-firm digital expertise are jointly relevant to audit quality (Cao et al., 2015; Gepp et al., 2018; Rahman & Ziru, 2023).

Auditor–client networks and regulatory inspection outcomes

Network analysis offers a natural representation of an audit firm's public-client portfolio. A bipartite network could describe connections among firms, engagement partners, issuers, and component auditors. Theoretical network research distinguishes simple size from concentration. centrality. and diversity. All of which may influence coordination. information flow (Borgatti&Halgin. 2011; Newman. 2010; Wasserman & Faust. 1994).In the audit setting. PCAOB Form AP is especially useful since it identifies the filing audit firm. issuer CIK. engagement partner. report date. and certain participating firms.At the same time. inspection outcomes must not be overgeneralised: Part I. Findings reflect selected audits. reviewed areas. not an audit firm's complete population. The appropriate outcome is thus a firm-year inspection-risk measure, not a claim that every client of a cited firm was deficient.

Recent evidence

Table 1. Eight recent studies informing the research design.

Study	Setting	Relevant finding
Fedyket al.(2022)	U.S.audit firms; archival AI investment	AI investment was associated with improved audit quality and lower fees; evidence is firm-level, not filing-level.
Huang et al.(2022)	Conceptual/technical audit analytics	Proposes full-population testing through audit-by-exception; stresses implementation and judgment challenges.
Rahman &Ziru(2023)	Chinese listed firms	Client digitalisation and audit-firm digital expertise jointly relate to audit quality.
Föhret al.(2026)	Risk-based auditing framework	Develops a framework for deploying foundation and task-specific models while maintaining audit relevance.
Murikahet al.(2024)	Systematic review	Identifies bias, transparency, accountability, and fairness as central risks of AI in auditing.
Li & Goel (2025)	AI-auditability interviews/review	Classifies auditability measures and competencies needed to audit AI-driven systems.
Kokinaet al.(2025)	Field interviews with audit professionals	Finds substantial practical opportunities and constraints in AI adoption across audit work.
DeFond (2025)	Analytical review of audit failures	Highlights the need to distinguish observed deficiencies, systemic causes, and audit-quality consequences.

Table 1 shows that recent studies support the potential value of data analytics and artificial intelligence but also reveal a gap. Most studies focus on adoption, professional perceptions, broad audit quality proxies, or conceptual frameworks. Few combine (a) transparent XBRL and filing-text measures, (b) public auditor–client network data, (c) the regulatory Part I.A outcome, and (d) explainable, temporally validated machine learning. The present study addresses that gap with the following research questions:

RQ1: Is a higher auditor-portfolio DDCI associated with a higher subsequent Part I.A deficiency risk?

RQ2: Do portfolio concentration, partner workload, and component-auditor participation improve prediction beyond disclosure complexity?

RQ3: Which factors consistently explain predicted risk after calibration and temporal validation?

Theoretical mechanism and testable propositions

The theoretical mechanism is an information-processing and evidence-reliability argument. A filing with a larger, more heterogeneous, and more difficult-to-interpret digital footprint may increase the cognitive and technological demands placed on the audit team. Complexity could arise from genuine business conditions, including acquisitions, multi-segment operations, complex revenue arrangements, fair-value estimates, and new reporting requirements. It may also arise from data-management choices, such as extensive company-specific tagging or inconsistent narrative explanations. Neither source automatically implies low audit quality. The prediction is conditional: when complexity is concentrated within an auditor's portfolio and is not matched by sufficient coordination capacity, the probability of a subsequent regulator-observed deficiency may increase.

Auditor–client networks provide the portfolio-level boundary conditions for that mechanism. A diversified portfolio might generate knowledge spillovers across clients and industries; conversely, excessive portfolio concentration, unusually high partner workloads, or extensive component-auditor participation might increase coordination and review demands. In a firm-year panel, the relevant quantity is therefore not the complexity of one issuer in isolation but the exposure created by the complete observable portfolio of issuer clients. This design also avoids attributing a regulatory finding to an individual client when PCAOB public disclosures do not identify that issuer.

The empirical propositions are stated as predictive rather than causal hypotheses. P1 predicts that a higher portfolio-weighted DDCI at time t is associated with a higher probability of high Part I.A deficiency risk at t + 1. P2 predicts that network capacity constraints—higher concentration, partner workload, and component-auditor participation—

improve prediction beyond reporting-complexity features. P3 predicts that explanatory patterns will be considered substantively credible only when their SHAP direction is stable across temporal test folds and their predicted probabilities are well calibrated. These propositions discipline the subsequent machine-learning analysis by requiring temporal precedence, stable explanations, and auditable decision rules rather than relying on model fit alone.

Materials and Methods

Research design and unit of analysis

The full study adopts a longitudinal, firm-year predictive design. The unit of analysis is an inspected audit-firm-year, as the public Part I. An inspection file provides an audit-firm outcome but does not identify inspected issuers by CIK. SEC issuer features are therefore first calculated at the issuer-year level and subsequently aggregated across the lead auditor's Form AP client portfolio. The design is prospective: features are measured from SEC filings and Form AP records available at t, while the PCAOB inspection outcome is measured from the subsequent inspection report available at t + 1. This ordering eliminates the use of post-inspection disclosures and reduces look-ahead bias.

Data sources and linkage rules

Table 2.Public data sources and the role of each source in the completed panel.

Source	Core fields	Use in study
SEC EDGAR/XBRL	Issuer financial facts, taxonomy tags, filing metadata, and narrative 10-K text	Construct issuer-year DDCI; retain only filings publicly available before the outcome period.
PCAOB Form AP /AuditorSearch	Firm ID, issuer CIK, report date, fiscal period end, engagement partner, and participating firms	Create a time-stamped auditor–client graph and portfolio/network measures.
PCAOB Firm Inspection Reports / Part I.A	Firm identity, inspection year, total audits reviewed, Part I.A deficiencies and detailed findings	Construct firm-year deficiency-rate outcome and conduct the public-data pilot.

SEC APIs provide comparable XBRL facts through company-concept, company-facts, and frame endpoints, while bulk archives support large-scale extraction (SEC, 2025). Form AP reports are filed for public-company audit reports and publicly disclose the key audit participants necessary for network construction (PCAOB, 2026a). The linkage key is the issuer CIK and fiscal-period end date, matched to the Form AP audit report date within a prespecified filing window. The PCAOB inspection file is then joined by audit-firm registration identifier and inspection year. All amendments are deduplicated by retaining the latest Form AP filing. Client observations with missing lead-firm or fiscal-period information are excluded before aggregation, and no inference is made from nonpublic engagement details.

Variable operationalisation

Digital disclosure complexity is measured as a standardized composite of six components. Let z(·) denote a year-standardized score. For issuer i in fiscal year t, the index is defined as follows:

$$DDCI_{it} = \frac{\left(z(ExtensionRatio_{it}) + z(TagDensity_{it}) + z(NarrativeVolume_{it}) + z(ReadabilityDifficulty_{it}) + z(RiskDensity_{it}) + z(NumericNarrativeGap_{it}) \right)}{6} \quad (1)$$

Extension Ratio is the number of entity-specific XBRL tags divided by the total number of tags; Tag Density is the number of distinct tags per reported fact; Narrative Volume is the logarithm of the 10-K word count; Readability Difficulty is the Gunning Fog Index; Risk Density is the number of risk-related words divided by the total word count; and Numeric–Narrative Gap is the absolute standardized difference between numeric deterioration signals and textual risk intensity. The firm-level portfolio measure is the client-weighted mean of DDCI, supplemented by the 90th percentile and within-portfolio standard deviation to identify concentrations of highly complex clients.

The auditor–client graph is defined as G_t = (F_t, I_t, E_t), where F_t is the set of audit firms, I_t is the set of issuer clients, and E_t represents the Form AP links at time t. Network measures include firm client degree, Herfindahl concentration, industry entropy, partner workload, and component-auditor participation. For an audit firm f, client concentration is defined as follows:

$$HHI^{f}_{t} = \sum_{i} \left(\frac{w_i f_t}{\sum_i w_i f_t} \right)^2 \quad (2)$$

where w is equal-weighted client exposure or, in robustness tests, client market-capitalisationweight. The principal outcome is a firm-year Part I. A deficiency rate: the number of inspected audits with at least one Part I.A deficiency divided by total selected audits reported by the PCAOB.A high-risk binary target is created only after a preregistered threshold is chosen from the distribution of completed historical data; the preferred specification uses the annual upper quartile to avoid importing an arbitrary regulatory benchmark.

Table 3. Operational data dictionary for the completed firm-year predictive panel.

Construct	Operational definition	Source
Portfolio DDCI	Client-weighted mean of issuer DDCI components	SEC XBRL and 10-K text
Upper-tail DDCI	90th percentile client DDCI within firm-year	SEC XBRL and 10-K text
Disclosure dispersion	Standard deviation of client DDCI within firm-year	SEC XBRL and 10-K text
Client degree	Number of issuer clients linked to the audit firm	PCAOB Form AP
Partner workload	Number of Form AP issuer reports per engagement partner	PCAOB Form AP
Client concentration	Herfindahl index of client exposure	PCAOB Form AP and SEC market cap
Industry diversity	Shannon entropy of client SIC/NAICS groups	Form AP and SEC filings
Component involvement	Share of auditor-client links reporting other audit firms	PCAOB Form AP
Part I.A rate	Affected inspected audits divided by selected audits reviewed	PCAOB reports and summary data
High-risk target	Upper-quartile deficiency rate in historical training years	Derived within training windows

Machine-learning models and key equations

Four models are estimated. Logistic regression is the transparent benchmark. Elastic-net logistic regression handles correlated disclosure and network features while preserving sparse selection. Random forest captures nonlinear interactions without a functional-form assumption, and XGBoost provides boosted tree predictions. The logistic model is:

$$Pr(Yf_{,t+1} = 1 \mid Xf_t) = \frac{1}{1 + \exp\left(-(\beta^0 + Xf_t\beta)\right)} \quad (3)$$

The elastic-net estimator minimizes the negative log-likelihood plus a combined L1/L2 penalty:

$$L(\beta) = -\ell(\beta) + \lambda \left[\alpha \left|\left|\beta\right|\right|^1 + \frac{(1 - \alpha)\left|\left|\beta\right|\right|^{22}}{2} \right] \quad (4)$$

For XGBoost, the objective at iteration m is the regularised sum of losses plus tree complexity:

$$Obj(m) = \sum l\left(y_i, \hat{y}_i(m-1) + f_m(x_i)\right) + \Omega(f_m) \quad (5)$$

Hyperparameters are tuned only inside the training window.SHAP explanations are used only after the final model is selected; the feature contributionφjis the average marginal contribution of feature j over all feature coalitions. This supports local explanation of a firm-year prediction and global ranking of stable drivers (Lundberg & Lee, 2017).

Validation, calibration, and quality criteria

The design uses expanding-window temporal validation. For example, models trained on 2019–2021 are tuned within 2019–2021 and evaluated on 2022; the window then expands

forward. The final test year is held out once. Imputation, winsorisation, standardisation, text-vector construction, feature selection, resampling, and model tuning occur within each training window only. This separation is necessary as randomly shuffling firm-years would leak temporal information (Varma & Simon, 2006; Roberts et al., 2017). Class imbalance is addressed by class-weighted learning inside the training data rather than by synthetic observations in the test year.

Table 4. Model-evaluation criteria and required safeguards

Criterion	Measures	Decision rule
Discrimination	PR-AUC (primary), ROC-AUC, balanced accuracy, F1	PR-AUC is emphasised when high-risk observations are relatively uncommon.
Calibration	Brier score, calibration intercept/slope, reliability plot	Predicted probabilities must agree with observed event frequencies.
Stability	Fold-to-fold range; bootstrap confidence intervals	A high mean score with unstable folds is not accepted as reliable.
Explainability	SHAP direction, dependence plots, and sign consistency	Explanations are descriptive, not causal effects.
Audit relevance	Temporal availability and no post-outcome variables	Every feature must be observable before inspection reporting.

6. Sensitivity analysis, reproducibility, and governance

Three sensitivity analyses are pre-specified. First, DDCI is estimated both as an equal-weighted composite and by principal-component weights learned only in the training window. Second, portfolio exposure is calculated with equal client weights and with lagged market-capitalisationweights. Third, the outcome is estimated as both an annual upper-quartile binary risk indicator and the continuous Part I. A deficiency rate. The study reports whether substantive drivers remain directionally consistent across these specifications, rather than selecting the specification with the most favourable score.

The replication package needs to retain raw-file URLs. retrieval dates. version hashes. data-cleaning rules. and an audit trail for every CIK-to-Form-AP-to-firm-year merge. The code must record the availability date of every 10-K, Form AP, and PCAOB report to prevent look-ahead bias. All feature engineering, including text preprocessing and standardisation, is re-estimated inside the training partition. This governance layer is necessary as an explainable model is not auditable when its inputs, transformations, or versions cannot be reproduced.

Results

Public-data audit and sample description

The empirical pilot used the PCAOB Part I.A downloadable dataset obtained on 7 July 2026. The raw file contains 13,743 finding records from 391 registered audit firms and inspection

years 2018–2025. Because 2025 contains only 44 affected inspected-audit observations in this snapshot, the principal descriptive window is restricted to 2018–2024. The resulting complete-window data contain 1,698 affected inspected-audit observations. This figure does not represent 1,698 unique public issuers: the PCAOB uses an anonymised issuer reference key that identifies an observation only within an inspection report. The unit is consequently correctly described as an affected inspected-audit observation, not a cross-year client panel.

Table 5. Descriptive PCAOB Part I.A. Data for complete inspection years, 2018–2024. High severity is an analytical proxy defined as four or more cited Part I.A findings within an affected inspected audit

Year	Firms	Affected inspected-audit observations	Median Part I.A findings	Mean Part I.A findings	High-severity proxy
2018	10	85	3	3.65	41.2%
2019	109	260	3	3.77	37.7%
2020	79	181	2	3.21	31.5%
2021	91	232	3	3.64	38.8%
2022	101	284	3	4.50	45.8%
2023	119	361	3	4.63	47.1%
2024	104	295	3	4.34	42.7%

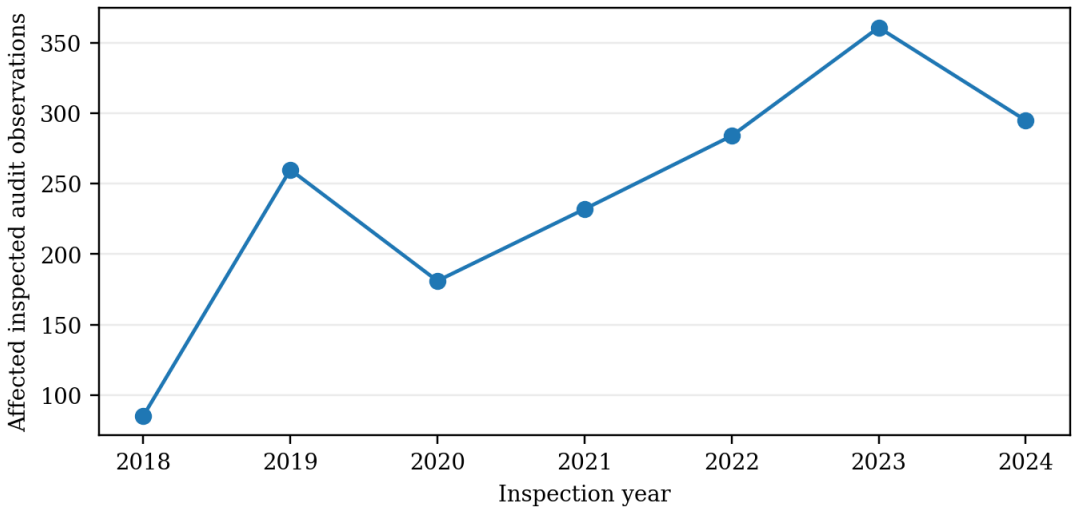


Figure 1. Affected inspected-audit observations in the public PCAOB Part I.A data, 2018–2024

The number of affected inspected-audit observations increased from 85 in 2018 to 361 in 2023, then declined to 295 in 2024. This is a description of the released finding data, not a population trend in all public-company audits. Changes might reflect inspection coverage, selection approaches, reporting timing, and the mix of firms and audit areas reviewed. The PCAOB explicitly notes that selected audits are not representative of a firm's whole population. That comparisons across years or firms require contextual caution (PCAOB. 2026b).

Severity pattern of disclosed Part I.A findings

Table 6. Classification of affected inspected audits in the 2018–2024 PCAOB Part I.A subset

PCAOB classification	Affected inspected-audit observations	Percentage
multiple deficiencies	1322	77.9%
a single deficiency	336	19.8%
an incorrect opinion on the financial statements and/or ICFR	40	2.4%

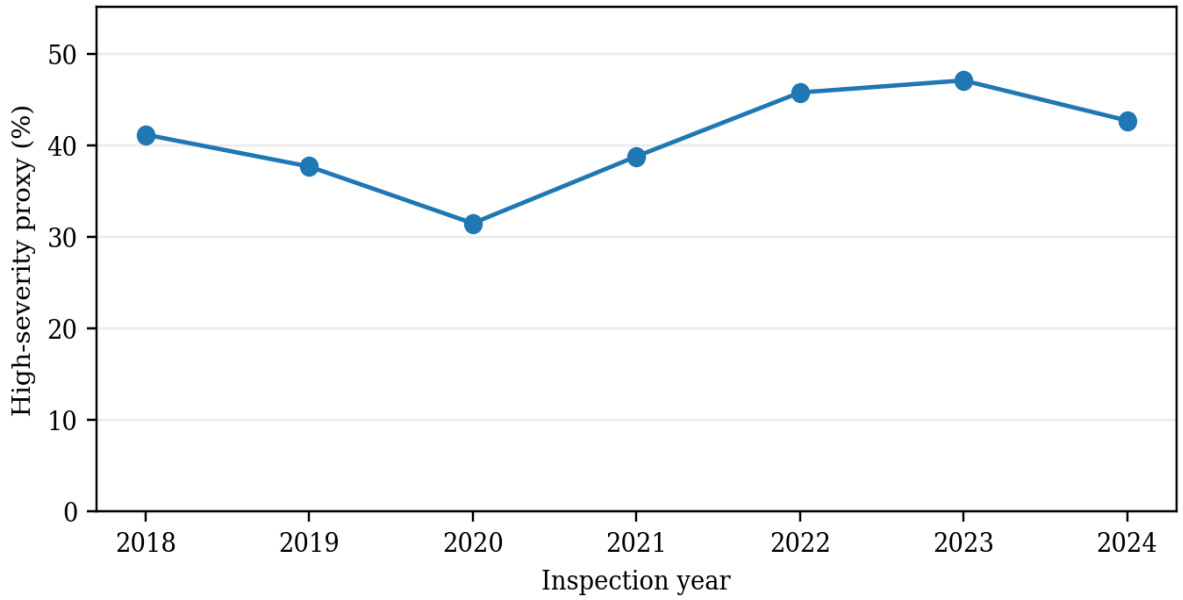


Figure 2.Annual prevalence of the analytical high-severity proxy in affected inspected audits, 2018–2024.

Of the 1,698 affected inspected-audit observations. 1,322 (77.9%) were classified by the PCAOB as audits with multiple deficiencies. 336 (19.8%) as audits with a single deficiency. and 40 (2.4%) as audits with an incorrect financial-statement and/or ICFR opinion. The high-severity proxy ranged from 31.5% in 2020 to 47.1% in 2023. These results do not measure the probability that any randomly selected audit will be deficient; Part I.A file contains cited findings rather than the complete set of reviewed non-deficient audits. Its value for the present research is as a regulator-grounded outcome source after it is combined with the inspection-summary file. which supplies the total audits reviewed.

https://jitm.ut.ac.ir/

Audit areas in the disclosed records

Table 7. Ten most frequent audit-area labels in U.S.annually inspected firms, 2018–2024. Counts are coded finding-line occurrences, not unique engagement counts

Audit area (PCAOB label)	Coded finding-line occurrences
Revenue	1161
Inventory	392
Business Combinations	205
Accounts Receivable	203
Business Combination	165
Deferred Revenue	131
Allowance for Loan Losses	109
Investment Securities	94
Allowance for Credit Losses	88
Goodwill	87

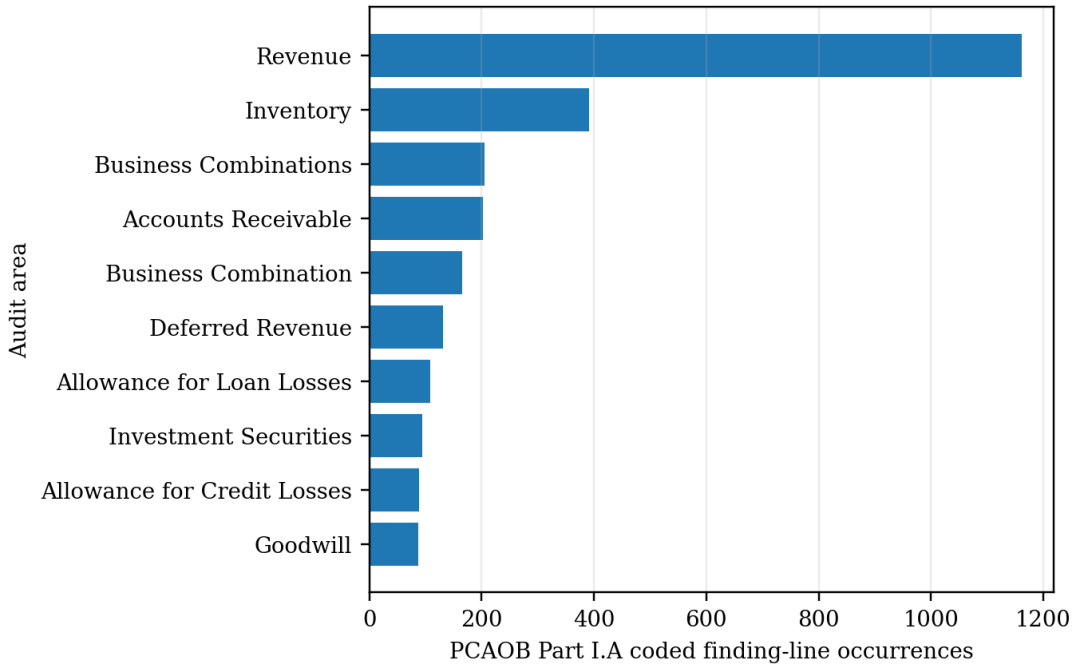


Figure 3. Most frequent PCAOB-coded audit areas among U.S.annually inspected firms, 2018–2024.

Revenue is the most frequently coded area in the focused display, followed by inventory, business combinations, accounts receivable, deferred revenue, and credit-loss-related areas. This pattern is compatible with the audit significance of revenue recognition. valuation. estimation. and transaction-specific judgments; nevertheless. it is not evidence that an audit area is intrinsically deficient. Inspection teams select areas based on risk and may focus on emerging or complex topics. Accordingly, audit-area frequencies are used to motivate the content dimensions of DDCI and not as labels for a machine-learning model.

Implications for the completed SEC–Form AP–PCAOB model

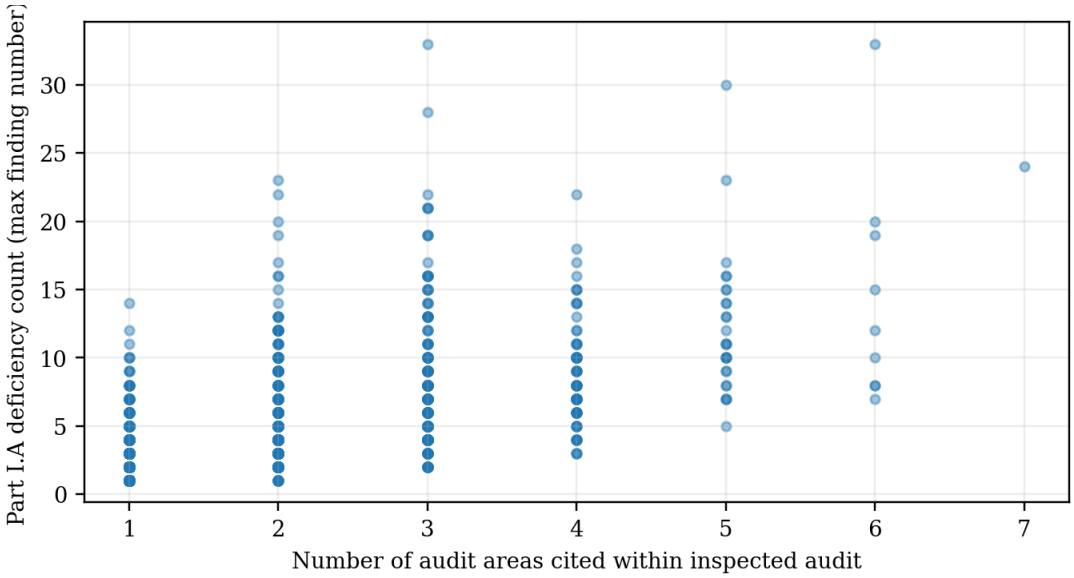


Figure 4. Exploratory association between cited audit-area breadth and the number of Part I.A findings in affected inspected audits, 2018–2024.

The exploratory Spearman correlation between cited audit-area breadth and the number of Part I.A findings is 0.680 ($p < .001$). This association is expected, as both quantities are measured after inspection based on the cited findings; therefore, it is not used as a predictor and must not be interpreted as a causal result. The completed model instead uses disclosure and network measures available before the inspection outcome. Specifically, it will estimate whether a firm's client portfolio has a higher subsequent deficiency risk when its average DDCI is higher, after controlling for firm size, portfolio size, country exposure, industry composition, inspection type, and year effects.

No AUC, F1-score, or SHAP ranking is reported for the completed predictive model in this pilot. The public Part I.A detailed file does not reveal issuer CIKs, and producing issuer-level model performance without a valid Form AP link would constitute an invalid linkage. This is a substantive result of the data audit: a Q1-quality design must not substitute anonymous inspection keys for SEC issuer identifiers. Once the Form AP client graph and inspection-summary denominator are merged, the specification in Section 3 permits a firm-year model with a defensible outcome, temporal ordering, and transparent validation. The present findings thus establish the regulatory outcome landscape and demonstrate why revenue-related and valuation-related digital disclosures ought to be retained as candidate dimensions rather than treated as ex post predictors.

Interpretation and boundary conditions

The descriptive evidence ought to be interpreted as regulator-published cited findings, not as a full distribution of all audits or as a scorecard of audit firms. Inspection selection, report

timing, changes in standards, and the mix of industries examined may affect the observed counts. The increase in affected audit observations through 2023 and the concentration of cited areas around revenue and valuation-sensitive accounts are consequently useful for defining model-relevant disclosure dimensions, but they cannot support a causal conclusion that digital complexity produces deficiencies. The full panel must include year effects, inspection-type indicators, firm size, country exposure, and industry composition, and it must separate feature construction from the future report date.

For journal submission, the predictive stage must be completed only after the inspection-summary denominator and Form AP portfolio link have been independently validated. The minimum reporting standard is a held-out final-year assessment with PR-AUC, calibration slope, Brier score, a confusion matrix at a preregistered operating threshold, and fold-level SHAP stability. A model that is accurate but poorly calibrated, temporally unstable, or driven by unavailable post-outcome information must not be characterized as an audit-quality tool. These safeguards turn the study from a generic machine-learning application into a defensible audit-analytics design.

Conclusion

This study develops a rigorous public-data framework for assessing whether digital disclosure complexity and auditor–client network structure are associated with regulator-observed audit-deficiency risk. Its central contribution is to align the level of analysis with what public sources genuinely identify. SEC EDGAR/XBRL supports issuer-level measurement of structured and narrative reporting complexity. PCAOB Form AP supports a time-stamped auditor–client network. PCAOB inspection data provide a regulatory outcome, but the detailed Part I.A file uses anonymized issuer reference keys; it cannot be treated as an issuer-level label without the additional Form AP and inspection-summary linkage. The study therefore treats this constraint as a methodological result rather than manufacturing predictions from incompatible identifiers.

The public-data audit identifies 13,743 Part I.A finding records 1,742 affected inspected-audit observations and 391 firms in the available 2018–2025 snapshot. In the complete 2018–2024 window, most affected audits are classified as involving multiple deficiencies. Revenue, inventory, receivables, business combinations, and valuation-related areas occur frequently in cited records. These patterns justify including both structured-reporting and narrative-complexity dimensions in the future portfolio DDCI, but they do not establish causal effects, rates across all audits, or indicate that every audit involving these areas is deficient.

The completed empirical study must aggregate issuer-level DDCI measures across each Form AP client portfolio, predict a subsequent firm-year Part I.A deficiency-rate outcome, and evaluate models using expanding-window validation, calibration, and SHAP stability. A

useful contribution to audit research requires more than a high discrimination score: features must predate the outcome, predicted probabilities must be calibrated, and explanations must be stable across years. This design combines transparent public data with regulator-grounded evidence while preserving the essential distinction between prediction, explanation, and overall audit-quality assessment.

Declaration of Competing Interest

The authors declare that they hold no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Data Availability Statement

The raw data used in this study are publicly available from official regulatory sources. Financial filing and XBRL data were obtained from the U.S. Securities and Exchange Commission (SEC) EDGAR database. While audit inspection. Auditor-reporting information was obtained from the Public Company Accounting Oversight Board (PCAOB) public repositories. including firm inspection reports. Form AP/AuditorSearchrecords.

The data-construction procedures. Variable definitions. and model-estimation scripts used to produce the empirical results will be made available in an open research repository upon acceptance of the article. No proprietary, confidential, or personally identifiable data were used in this study.

CRediT Authorship Contribution Statement

NazarMaanAbdul Kareem: Conceptualization, Methodology, Software, Formal Analysis, Data Curation, Visualization, Writing – Original Draft Preparation, and Writing – Review & Editing.

Ammar Ghazi Ibrahim Al-Ezzi: Conceptualization, Methodology, Validation, Supervision, and Writing – Review & Editing.

Waleed Khamees Sabea: Investigation, Validation, Data Curation, and Writing – Review & Editing.

All authors have read and approved the final version of the manuscript.

Acknowledgements

The authors acknowledge the U.S. Securities and Exchange Commission and the Public Company Accounting Oversight Board for maintaining publicly accessible regulatory data resources that enabled this study. The authors are grateful to their respective academic institutions for providing an environment that supported the development of this research.

The interpretations. analyses. and conclusions expressed in this article are solely those of the authors. do not necessarily represent the views of the SEC. The PCAOB or any affiliated institution.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

The author's received no financial support for the research, authorship, and/or publication of this article.

References

Aobdia, D. (2018). The impact of the PCAOB individual engagement inspection process—Preliminary evidence. *The Accounting Review, 93*(4), 53–80. https://doi.org/10.2308/accr-51911

Aobdia, D. (2019). Do practitioner assessments agree with academic proxies for audit quality? Evidence from PCAOB and internal inspections. *Journal of Accounting and Economics, 67*(1), 144–174. https://doi.org/10.1016/j.jacceco.2018.11.004

Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big Data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory, 36*(4), 1–27. https://doi.org/10.2308/ajpt-51684

Bao, Y., Ke, B., Li, B., Yu, J., & Zhang, J. (2020). Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. *Journal of Accounting Research, 58*(1), 199–235. https://doi.org/10.1111/1475-679X.12292

Bierstaker, J. L., Burnaby, P., & Thibodeau, J. (2001). The impact of information technology on the audit process: An assessment of the state of the art and implications for the future. *Managerial Auditing Journal, 16*(3), 159–164. https://doi.org/10.1108/02686900110385489

Borgatti, S. P., & Halgin, D. S. (2011). On network theory. *Organization Science, 22*(5), 1168–1181. https://doi.org/10.1287/orsc.1100.0641

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review, 78*(1), 1–3. https://doi.org/10.1175/1520-0493(1950)078<0001>2.0.CO;2

Brown-Liburd, H., Issa, H., & Lombardi, D. (2015). Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions. *Accounting Horizons, 29*(2), 451–468. https://doi.org/10.2308/acch-51023

Cao, M., Chychyla, R., & Stewart, T. (2015). Big Data analytics in financial statement audits. *Accounting Horizons, 29*(2), 423–429. https://doi.org/10.2308/acch-51068

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Davis, J., & Goadrich, M. (2006). The relationship between precision-recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning*, 233–240. https://doi.org/10.1145/1143844.1143874

DeFond, M. (2025). Audit failures: Why they occur and some suggestions for dealing with them. *Accounting and Business Research*. Advance online publication. https://doi.org/10.1080/00014788.2025.2516269

DeFond, M., & Zhang, J. (2014). A review of archival auditing research. *Journal of Accounting and Economics, 58*(2–3), 275–326. https://doi.org/10.1016/j.jacceco.2014.09.002

Earley, C. E. (2015). Data analytics in auditing: Opportunities and challenges. *Business Horizons, 58*(5), 493–500. https://doi.org/10.1016/j.bushor.2015.05.002

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fedyk, A., Hodson, J., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies, 27*(3), 938–985. https://doi.org/10.1007/s11142-022-09697-x

Financial Reporting Council. (2025, June 26). *AI in audit: Illustrative example and documentation guidance*. https://www.frc.org.uk/library/standards-codes-policy/audit-assurance-and-ethics/guidance/ai-in-audit/

Föhr, T. L., Schreyer, M., Moffitt, K. C., & Marten, K.-U. (2026). Deep learning meets risk-based auditing: A holistic framework for leveraging foundation and task-specific models in audit procedures. *International Journal of Accounting Information Systems, 57*, Article 100758. https://doi.org/10.1016/j.accinf.2025.100758

Francis, J. R. (2011). A framework for understanding and researching audit quality. *Auditing: A Journal of Practice & Theory, 30*(2), 125–152. https://doi.org/10.2308/ajpt-50006

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gao, R., Huang, S., & Wang, R. (2021). Data analytics and audit quality. *Singapore Management University School of Accountancy Research Paper No. 2022-151*. https://doi.org/10.2139/ssrn.3928355

Gepp, A., Linnenluecke, M. K., O'Neill, T. J., & Smith, T. (2018). Big data techniques in auditing research and practice: Current trends and future opportunities. *Journal of Accounting Literature, 40*, 102–115. https://doi.org/10.1016/j.acclit.2017.05.003

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering, 21*(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239

Huang, F., No, W. G., Vasarhelyi, M. A., & Yan, Z. (2022). Audit data analytics, machine learning, and full population testing. *Journal of Finance and Data Science, 8*, 138–144. https://doi.org/10.1016/j.jfds.2022.05.002

Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting, 13*(2), 1–20. https://doi.org/10.2308/jeta-10570

Kokina, J., & Davenport, T. H. (2017). The emergence of artificial intelligence: How automation is changing auditing. *Journal of Emerging Technologies in Accounting, 14*(1), 115–122. https://doi.org/10.2308/jeta-51730

Kokina, J., Blanchette, S., Davenport, T. H., & Pachamanova, D. (2025). Challenges and opportunities for artificial intelligence in auditing: Evidence from the field. *International Journal of Accounting Information Systems, 56*, Article 100734. https://doi.org/10.1016/j.accinf.2025.100734

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Li, Y., & Goel, S. (2025). Artificial intelligence auditability and auditor readiness for auditing artificial intelligence systems. *International Journal of Accounting Information Systems, 56*, Article 100739. https://doi.org/10.1016/j.accinf.2025.100739

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

Moffitt, K. C., Rozario, A. M., & Vasarhelyi, M. A. (2018). Robotic process automation for auditing. *Journal of Emerging Technologies in Accounting, 15*(1), 1–10. https://doi.org/10.2308/jeta-10589

Murikah, W., Nthenge, J. K., & Musyoka, F. M. (2024). Bias and ethics of AI systems applied in auditing—A systematic review. *Scientific African, 25*, e02281. https://doi.org/10.1016/j.sciaf.2024.e02281

Newman, M. (2010). *Networks: An introduction*. Oxford University Press.

Public Company Accounting Oversight Board. (2024). *Amendments related to aspects of designing and performing audit procedures that involve technology-assisted analysis of information in electronic form.* https://pcaobus.org/oversight/standards/standard-setting-research-projects/amendments-related-to-certain-aspects-of-designing-and-performing-audit-procedures-that-involve-technology-assisted-data-analysis

Public Company Accounting Oversight Board. (2025a, April 28). *PCAOB makes available new downloadable datasets featuring PCAOB inspection findings from audit firm inspection reports*. https://pcaobus.org/news-events/news-releases/news-release-detail/pcaob-makes-available-new-downloadable-datasets-featuring-pcaob-inspection-findings-from-audit-firm-inspection-reports

Public Company Accounting Oversight Board. (2025b, October 1). *Staff guidance—Examples of evaluating the relevance and reliability of audit evidence obtained from external sources*. https://pcaobus.org/oversight/standards/implementation-resources-PCAOB-standards-rules/amendments-related-to-aspects-of-designing-and-performing-audit-procedures-that-involve-technology-assisted-analysis-of-information-in-electronic-form

Public Company Accounting Oversight Board. (2025c). *AS 1105: Audit evidence*. https://pcaobus.org/oversight/standards/auditing-standards/details/AS1105

Public Company Accounting Oversight Board. (2026a). *AuditorSearch*. https://pcaobus.org/resources/auditorsearch

Public Company Accounting Oversight Board. (2026b). *Firm inspection reports*. https://pcaobus.org/oversight/inspections/firm-inspection-reports

Rahman, M. J., & Ziru, A. (2023). Clients' digitalization, audit firms' digital expertise, and audit quality: Evidence from China. *International Journal of Accounting & Information Management, 31*(2), 221–246. https://doi.org/10.1108/IJAIM-08-2022-0170

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., & Hartig, F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography, 40*(8), 913–929. https://doi.org/10.1111/ecog.02881

Securities and Exchange Commission. (2025, April 8). *EDGAR application programming interfaces (APIs)*. https://www.sec.gov/search-filings/edgar-application-programming-interfaces

Steyerberg, E. W. (2019). *Clinical prediction models* (2nd ed.). Springer. https://doi.org/10.1007/978-3-030-16399-0

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B, 58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics, 7*, Article 91. https://doi.org/10.1186/1471-2105-7-91

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big Data in accounting: An overview. *Accounting Horizons*, *29*(2), 381–396. https://doi.org/10.2308/acch-51071

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815478

XBRL International. (2021). *XBRL specification 2.1 with corrigenda corrections*. https://www.xbrl.org/specification/xbrl-2-1/

Appendix A. Reproducible Data-Assembly Protocol

This appendix specifies the operational sequence required to reproduce the completed firm-year panel. The purpose is to ensure that the public-data design remains auditable from raw retrieval through model evaluation. Each source must be downloaded with its retrieval date and content hash, and the original files must be retained without overwriting later versions. The panel is built at the auditor-firm-year level only after issuer-level SEC measures have been calculated and aggregated through Form AP.No variable may be calculated from a filing, report, or amendment that became public after the prediction date.

Appendix Table A1.Minimum reproducibility record for the completed SEC-Form AP-PCAOB panel

Stage	Required operation	Retained audit trail
1. SEC acquisition	Retrieve submissions, company facts, and filing text; retain 10-K forms and accession dates.	CIK, fiscal year, accession number, availability date
2.XBRL/text engineering	Calculate tags, extensions, word count, readability, risk density, and numeric-narrative gap.	Issuer-year DDCI components
3. Form AP graph	Deduplicate amended Form AP filings and create dated lead-auditor, issuer, partner, and component-firm links.	Firm-issuer-year network edges
4. Portfolio aggregation	Aggregate only clients observable at t to the audit-firm-year level; preserve counts and dispersion.	DDCI mean, tail, HHI, workload, diversity
5. PCAOB outcome	Extract inspection year, audits reviewed, and Part I.A affected audits; preserve inspection type and report date.	Firm-year deficiency rate and outcome date
6. Validation	Split by time before imputation, tuning, scaling, or text-vector construction; freeze final test year.	Leakage-safe evaluation records

https://jitm.ut.ac.ir/

Before model estimation, the researcher ought to reconcile audit-firm registration identifiers across PCAOB sources and document all unmatched CIK, fiscal-year, and report-date records. The unmatched set must not be silently discarded; its size, reason, and source distribution must be reported in a linkage-flow table. After aggregation, every feature ought to carry a data-availability timestamp and a clear definition of whether it is issuer-weighted, equally weighted, or market-capitalization-weighted. This protocol provides the documentation necessary for an independent reviewer to reproduce the data lineage.

The prediction design must also be assessed to ensure that it respects the chronological order of audit evidence.

Bibliographic information of this paper for citing:

Abdul Kareem, Nazar Maan; Ibrahim Al-Ezzi, Ammar Ghazi & Khamees Sabea, Waleed (2026). Explainable Machine Learning for Audit Risk Analytics: Digital Disclosure Complexity and Auditor–Client Network Structure. *Journal of Information Technology Management*, 18 (Special Issue), 102-121. https://doi.org/10.22059/jitm.2026.109975

Copyright © 2026, Nazar Maan Abdul Kareem, Ammar Ghazi Ibrahim Al-Ezzi and Waleed Khamees Sabea