

American Journal of Scholarly Research and Innovation

Volume: *14;* ***Issue:*** *12*
Pages: *41– 72*
eISSN: 3067-2163

Human-In-The-Loop Artificial Intelligence for Account Reconciliation and Financial-Reporting Transparency: A Bibliometric and Science-Mapping Analysis

Songeta Dhar[1];

[1]. Accounting & Business Analytics Researcher;
Master of Business Studies (MBS) in Accounting, National University, Bangladesh;
Bachelor of Commerce (Hons.) in Accounting, National University, Bangladesh;
Email: songeta.bd@gmail.com

Doi: 10.63125/jamadx91
Received: *19 September 2024;* ***Revised:*** *22 October 2024;* ***Accepted:*** *24 November 2024;* ***Published:*** *03 December 2024*

Abstract

Human-in-the-loop artificial intelligence (AI) is increasingly examined in accounting workflows involving transaction matching, exception detection, and financial close. This study aimed to map research connecting AI-assisted account reconciliation, human oversight, and financial-reporting transparency through bibliometric and science-mapping methods. The proposed search covered Scopus and Web of Science publications from 2015 to 2025. Records were screened for relevance, and bibliographic data were prepared for publication, citation, collaboration, and keyword analyses. In the screening example, 1,248 records were retrieved, 286 duplicates were removed, and 962 records were screened by title and abstract. After 261 full texts were assessed, 44 publications were included. The example corpus comprised 30 journal articles, eight conference papers, and six reviews; 36 publications, or 81.8%, were multi-authored. The dataset accumulated 612 citations, with a median of eight citations per publication. A count model estimated annual publication growth at an incidence rate ratio of 1.20, with a 95% confidence interval of 1.08 to 1.34 and a p-value of .001. Network mapping produced five author co-authorship clusters and four reference co-citation clusters. Keyword analysis retained 24 terms and identified five thematic groups addressing accounting automation, reconciliation and financial close, audit analytics and internal control, human oversight and explainability, and transparency and reporting quality. The comparison of citation distributions by document type was statistically significant before adjustment; after citation-age adjustment, it was not significant ($p = .086$). These example results demonstrate how bibliometric indicators can describe publication activity, citation patterns, collaboration, and conceptual organization. They do not establish that AI improves reconciliation accuracy, human decisions, or financial-reporting outcomes. Final conclusions require verified database records, screening decisions, citation data, and network analyses. The study's intended contribution is an integrated quantitative map of research at the intersection of AI-supported reconciliation, human review, and transparent financial reporting.

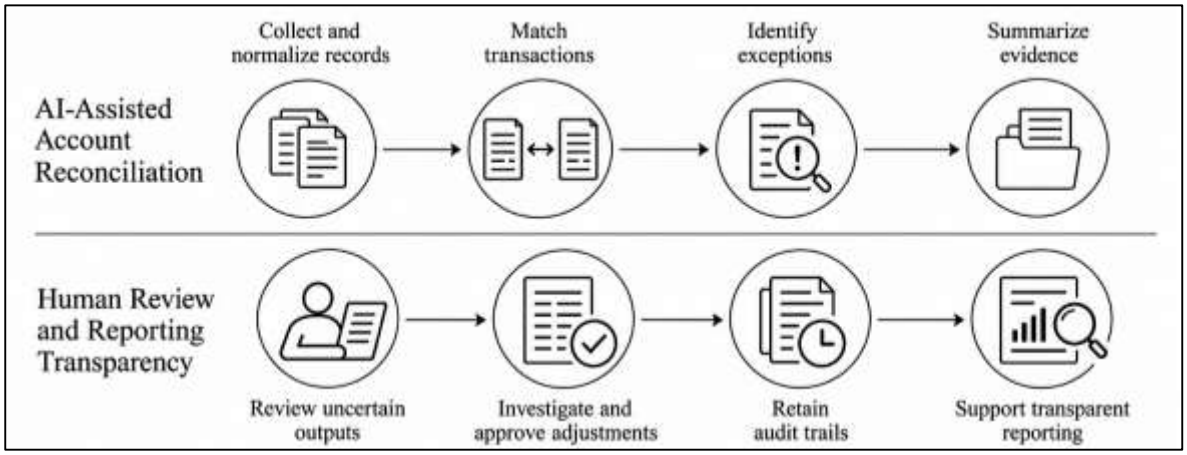
Keywords

Human Oversight; Artificial Intelligence; Account Reconciliation; Bibliometric Analysis; Financial Transparency.

INTRODUCTION

Account reconciliation is the process of comparing financial records that describe the same activity and resolving differences between them (Cao et al., 2021). Organizations reconcile bank statements with cash ledgers, subledger balances with general-ledger accounts, intercompany transactions across entities, and supporting schedules with reported balances. These procedures can confirm agreement, identify timing differences, and expose missing, duplicated, or incorrectly classified entries. Reconciliation connects transaction processing with financial-statement preparation by documenting how balances were checked, what exceptions arose, and how they were resolved. Financial-reporting transparency refers to the clarity, traceability, completeness, and understandability of information used to communicate an organization's financial position and performance. It depends on more than accurate totals: reviewers and other users need to understand how balances were derived, what evidence supports adjustments, and how decisions were controlled. Artificial intelligence (AI) encompasses computational methods that learn patterns, classify information, make predictions, process language, or support decisions. In accounting workflows, AI can extract data from documents, classify transactions, recommend matches, detect anomalies, prioritize exceptions, and summarize supporting evidence (Aksoy & Gurol, 2021).

Figure 1: Human Oversight in Account Reconciliation



Human-in-the-loop AI incorporates people into consequential stages of an AI-supported process. Accountants may define decision rules, review uncertain outputs, investigate exceptions, approve adjustments, or monitor system performance. The extent of human authority and review varies across systems, so a workflow should be understood by examining when people participate, what evidence they can access, and whether they can challenge or override recommendations. These concepts have international relevance because financial reporting spans multinational organizations, cross-border supply chains, global capital markets, and jurisdictions with different accounting and regulatory environments (Licht & Licht, 2020). Consolidation may require records from multiple currencies, systems, languages, and reporting calendars to be classified and reconciled consistently. Investors, lenders, regulators, employees, and suppliers use financial information to assess performance, obligations, stewardship, and risk. AI-enabled tools may help process large volumes of records, but their outputs must be interpreted within organizational policies and accounting contexts. A recommendation that appears plausible from transaction amounts or descriptions may not reflect the underlying economic event. Accordingly, research in this area concerns both technical performance and the review, documentation, and control of AI-supported decisions.

Reconciliation workflows typically involve collecting records, normalizing data, matching transactions, investigating unmatched items, documenting resolutions, and reviewing completed accounts. AI can support these stages through rule-based automation, machine-learning matching, natural-language processing, anomaly detection, and workflow routing (Licht & Licht, 2020). These functions should be distinguished: extracting information accurately does not establish that a reconciliation decision is

correct, and ranking a potential match does not confirm that two records represent the same accounting event. Research on accounting automation, machine learning, robotic process automation, audit analytics, anomaly detection, and intelligent document processing has examined related capabilities. Studies have measured outcomes such as processing time, match accuracy, exception rates, error detection, workload, and reporting quality, although methods and settings vary. Human oversight affects how these systems operate. Reviewers may confirm high-confidence matches, investigate uncertain recommendations, approve material adjustments, or sample completed transactions. Effective review depends on access to source evidence, visibility of uncertainty, the ability to revise outputs, clear escalation procedures, and retained decision records (Akter et al., 2022).

Research on human–automation interaction, algorithmic advice, professional judgment, explainable AI, and internal control offers relevant perspectives. However, a person's presence in a workflow does not itself establish meaningful oversight. Likewise, faster financial close or higher matching accuracy does not, by itself, demonstrate improved transparency. Transparency may involve traceability from source records to reported balances, understandable explanations, complete exception documentation, and accessible evidence of review. Bibliometric analysis and science mapping provide quantitative methods for describing publication activity and relationships among studies, authors, sources, references, and concepts. These techniques can identify influential works, collaboration patterns, citation structures, and thematic clusters across accounting, information systems, auditing, machine learning, and organizational research. Existing reviews have mapped AI and automation in accounting, AI in accounting and auditing, analytics and reporting quality, and related digital-accounting topics. Yet the specific intersection of human-in-the-loop AI, account reconciliation, and financial-reporting transparency remains distributed across these neighboring fields. Terminology also varies: studies may refer to financial close, bank reconciliation, exception detection, human–AI collaboration, auditability, or reporting quality (Hermann, 2022). This study quantitatively maps indexed scholarly literature concerning human participation in AI-assisted account reconciliation and its relationship to financial-reporting transparency. The analysis examines publication and collaboration patterns, influential sources and works, citation relationships, and keyword connections. Its conceptual scope includes AI-enabled matching, exception identification, document analysis, financial-close automation, human review, explainability, audit trails, internal control, reporting quality, and transparency. By distinguishing technical performance from accounting outcomes, and human presence from active oversight, the study provides a structured view of how these research areas are connected and organized.

The objective of this quantitative study is to map and characterize the scholarly literature on human-in-the-loop artificial intelligence for account reconciliation and financial-reporting transparency using bibliometric analysis and science-mapping techniques. The study aims to identify patterns in publication output over time, the journals and subject areas contributing to the topic, influential publications and references, and the authors, institutions, and countries represented in the indexed research. It also seeks to examine collaboration structures through co-authorship networks and to describe intellectual relationships through citation, co-citation, and bibliographic-coupling analyses. A further objective is to identify the main concepts and thematic groupings in the literature by analyzing keyword co-occurrence and related science maps, with attention to terms connected to AI, machine learning, reconciliation, financial close, human review, explainability, internal control, auditability, reporting quality, and transparency. The analysis will distinguish research focused on technical tasks, such as transaction matching, document extraction, classification, and anomaly detection, from research addressing organizational processes, including exception investigation, approval, accountability, and professional judgment. It will also examine how studies describe human participation in AI-supported workflows, including whether people define rules, review recommendations, resolve uncertain cases, approve adjustments, or monitor system outputs. In relation to financial reporting, the study will map how the literature operationalizes transparency and associated outcomes, including traceability, completeness of supporting evidence, reporting quality, and the documentation of decisions. Because publications may use different terminology, methods, settings, and levels of analysis, the study will also characterize how these differences appear across the mapped corpus and how they relate to the structure of the research field. Through these objectives, the

study provides a quantitative account of the volume, influence, collaboration, and conceptual organization of research at the intersection of human oversight, AI-enabled account reconciliation, and transparent financial reporting.

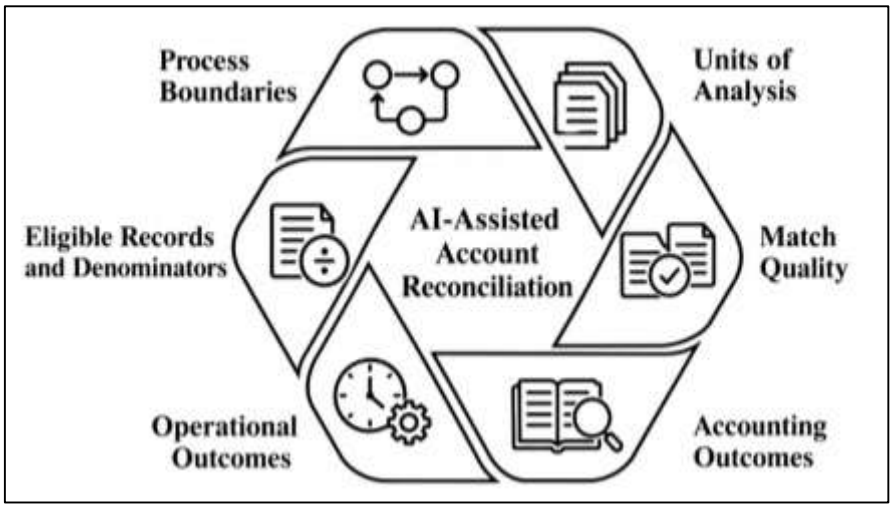
LITERATURE REVIEW

The literature on human-in-the-loop artificial intelligence (AI) for account reconciliation and financial-reporting transparency spans accounting information systems, financial close automation, machine learning, auditing, internal control, and human–AI collaboration. These areas address related but distinct questions. Technical studies assess how well systems extract, classify, match, or prioritize financial records; organizational studies examine how accountants use and review system recommendations; reporting research considers whether these processes support accurate, traceable, and understandable financial information. A literature review for this topic therefore needs to clarify the relevant concepts, identify how outcomes have been measured, and distinguish technical performance from accounting and transparency outcomes. This review is organized around quantifiable research dimensions that can also inform the bibliometric and science-mapping analysis. It examines the units of analysis, data sources, AI tasks, forms of human oversight, outcome measures, study designs, and research linkages used in the literature. Particular attention is given to how studies operationalize reconciliation accuracy, exception handling, reviewer intervention, financial-reporting quality, and transparency. The outline also considers how bibliometric indicators and network methods describe publication growth, influence, collaboration, and thematic structure. These distinctions provide a consistent framework for comparing research that uses different terminology, samples, and analytical methods.

AI-Assisted Account Reconciliation

Research on accounting automation and reconciliation treats reconciliation as a process for comparing records that should represent the same balance or economic event, identifying differences, and establishing how those differences were resolved (Aditya & Ahmed, 2022; Palaiogeorgou et al., 2021). The scope varies across studies. Some focus on bank reconciliation, where cash-book entries are compared with bank statements; others examine intercompany balances, subledger-to-general-ledger agreement, balance-sheet account substantiation, or transaction matching across invoices, purchase orders, payments, and ledger entries. Studies of machine-learning accounting services and financial-close automation often place reconciliation within a broader workflow that includes data extraction, classification, exception detection, and approval (Asheke & Maksudul, 2022; Drydakis, 2022). This distinction affects how performance is interpreted: a system may accurately match records but provide no support for investigation or close certification. Across research on robotic process automation, accounting information systems, document processing, anomaly detection, and audit analytics, process boundaries therefore range from a single matching task to the full reconciliation lifecycle. The lifecycle may include importing and standardizing records, generating candidate matches, assigning confidence or priority, investigating unmatched items, documenting explanations, proposing adjustments, obtaining approval, and confirming completion. Some studies evaluate only one stage, while others consider integration with financial close or control procedures (Asheke & Badrul, 2022; Konigsburg, 2022). Definitions also differ in whether timing differences count as resolved items, whether one-to-many matches are permitted, and whether a reconciliation requires documentary support or only agreement between system balances. These differences make it essential to identify exactly which stages a study includes when describing AI-assisted reconciliation. They also show why the presence of automation alone is insufficient to establish the completeness or quality of the reconciliation process. Studies use several units of analysis to assess reconciliation systems, and each unit captures a different level of performance. Transaction-pair studies evaluate whether individual records are correctly linked, while transaction-level datasets may include one-to-many or many-to-one matches (Zhang & Xin, 2019). Account-period studies assess whether a specific account has been reconciled for a defined reporting period. Exception-case studies focus on items requiring investigation, and user- or team-level studies examine workload, review behavior, or adoption. Organization-level analyses may assess close duration, control outcomes, or reporting quality over several accounting periods. These units are not directly interchangeable (Lefèvre et al., 2022; Akter & Sumaiya, 2022).

Figure 2: Operational Framework for AI Reconciliation



High transaction-level matching accuracy may coexist with unresolved material account balances, while a reduced number of exceptions may reflect improved matching or a change in how exceptions are classified. Denominators also vary across studies. Match rates may be calculated over eligible records, candidate pairs, or confirmed matches; exception rates may use transactions, accounts, or reconciliation cases; and review rates may refer to all AI recommendations or only items routed to a human (Imran & Jannatul, 2022; Silver, 2021). Completion measures may count accounts due for reconciliation, accounts submitted for review, or accounts formally certified. Research on accounting automation and machine-learning classification often reports record-level performance, whereas financial-close studies may report process-level time or completion outcomes. For meaningful comparison, a study needs to define its eligible population, exclusions, period, and denominator, including how duplicates, missing records, reversals, and unresolved items are handled. The literature consequently treats the unit and denominator as central parts of the operational definition rather than as minor reporting details. Clear specification allows readers to determine whether results describe model performance, reviewer workload, account completion, or organizational outcomes.

Research evaluating AI-assisted reconciliation uses a range of technical and operational indicators. Technical measures include the proportion of correctly identified matches, the proportion of true matches recovered, false matches, and records left unmatched (Maksudul & Islam, 2022; Tamburino & Bravo, 2021). Studies of machine-learning classification and document-processing systems may also assess performance through precision, recall, F1 scores, or related classification metrics. Operational measures include exception volume, the percentage of items requiring manual review, time to resolve exceptions, time spent per reconciliation, and total time to complete a close activity. Accounting-focused measures may include adjustment frequency, unresolved balance value, aged reconciling items, late certifications, and the proportion of reconciliations completed by deadline. These indicators reflect different consequences. A high match rate does not necessarily indicate low error if the system creates incorrect matches, and a reduction in manual review may not be beneficial if it results from missed exceptions. Likewise, faster completion is a process-efficiency outcome rather than direct evidence of improved reporting accuracy or transparency (Morshedul & Islam, 2022; Rungvivatjarus et al., 2020). Studies of automation in accounting services, bank-reconciliation support, anomaly detection, continuous auditing, and financial-close systems variously report transaction counts, percentages, monetary amounts, durations, and user workload. Their measures often depend on data availability and workflow design. For example, time savings may be assessed through system logs, observation, or user estimates, while unresolved balances may be assessed from account records at period end. Comparisons therefore require attention to how each indicator is defined, when it is measured, and whether it is evaluated against a manual, rule-based, or prior-period baseline. Taken together, the literature shows that a balanced assessment uses both match-quality measures and

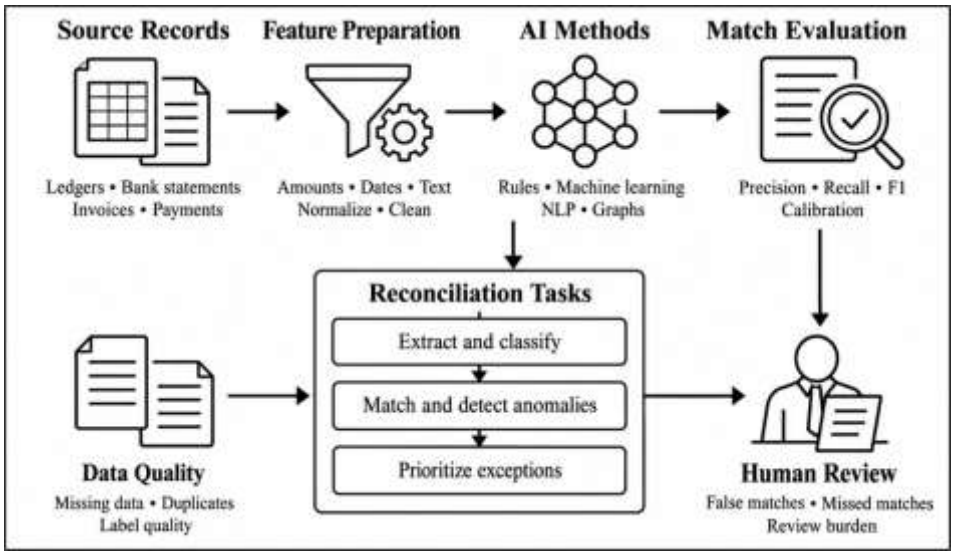
process or accounting outcomes, while preserving the distinction among them (Chiewchantanakit et al., 2020; Stabak & Shahid, 2022).

Across more than ten strands of research—including accounting information systems, robotic process automation, machine-learning classification, intelligent document processing, anomaly detection, bank-reconciliation support, financial-close automation, continuous auditing, audit analytics, and human–AI decision support—reconciliation is measured at different stages and scales. This body of work indicates that findings are most interpretable when process scope, unit of analysis, denominator, and outcome are reported together (Feroz & Pervaz, 2022; Tosun et al., 2022). For example, an evaluation of transaction matching should identify the records eligible for matching, how candidate pairs were generated, how many matches were manually confirmed, and how false matches were treated. An account-level evaluation should clarify which accounts and periods were included, how completion was defined, and whether unresolved balances or late adjustments were tracked. A human-review analysis should state whether the unit is a recommendation, exception, reviewer, or reconciliation session, and whether interventions include acceptance, correction, rejection, or escalation (Anyon et al., 2019; Feroz & Shabbir, 2022). These distinctions also apply to comparisons between AI-supported and manual procedures. Differences in transaction mix, account complexity, reporting deadlines, or data quality can affect observed match accuracy and completion time. When studies use inconsistent definitions, an apparent improvement in one measure may not correspond to improvement in another. A more complete description separates technical indicators, such as match errors, from operational indicators, such as review time, and accounting indicators, such as unresolved amounts or adjustments. It also records whether results are reported as counts, proportions, monetary values, or elapsed time. This operational structure provides a basis for synthesizing reconciliation research without treating unlike measures as equivalent, and it supports consistent classification of the studies included in the bibliometric and science-mapping analysis (Islam & Md, 2023; Jagannathan et al., 2020).

AI Methods and Technical Performance in Reconciliation Tasks

Studies of AI in accounting and reconciliation examine several distinct tasks: extracting information from financial documents, classifying transactions, matching records, detecting anomalies, and prioritizing exceptions for review. Rule-based systems apply predefined conditions, such as matching records with the same amount, date, or account identifier. They are useful when data are standardized and matching criteria are stable, but they can struggle when descriptions vary or records require contextual interpretation. Supervised machine-learning studies use previously labeled examples to classify transactions or rank candidate matches (Aksoy & Gurol, 2021; Asheke, 2023). Unsupervised methods identify unusual records or group transactions without relying on labeled outcomes, making them relevant to anomaly detection but potentially producing alerts that require substantial human review. Natural-language processing supports extraction and interpretation of text in invoices, bank descriptions, contracts, and correspondence. Graph-based approaches represent relationships among transactions, accounts, counterparties, and documents, which can help identify connected records when matching is not limited to a single pair (Filgueiras, 2022; Shahid, 2023). Generative AI has been studied for tasks such as summarizing evidence, explaining exceptions, and supporting interaction with accounting data; these uses differ from systems that make or confirm matches. Research on accounting automation, machine-learning accounting services, intelligent document processing, bank-reconciliation decision support, audit analytics, and anomaly detection thus encompasses a varied method set. The studies also differ in the level at which AI operates: some automate a narrow data-processing step, while others integrate model outputs into a broader workflow (Buchholtz, 2019; Islam & Ahmed, 2023). For synthesis, it is important to classify a method by its task and decision role, rather than by the general label "AI." A system that extracts a value, a model that ranks potential matches, and a tool that drafts an explanation address different technical problems and carry different risks of error.

Figure 3: AI Methods for Reconciliation Performance



AI performance in reconciliation depends on the source records and the way information is represented for analysis. Studies have used ledger entries, bank statements, invoices, purchase orders, payment records, account codes, and historical reconciliation outcomes. Common features include transaction amount, posting and value dates, currency, account identifier, counterparty, reference number, transaction description, document text, and the prior match history of similar records (Hermann, 2022; Sadia & Mosheur, 2023). Workflow data may add information about preparer, reviewer, approval status, exception category, or time to resolution. Data extraction studies focus on converting semi-structured documents into fields that can be compared, whereas matching studies combine structured fields and text-derived features to identify candidate relationships. Preparation steps may include currency conversion, date standardization, text cleaning, entity-name normalization, and harmonization of account codes across systems. Research on accounting data analytics and machine-learning classification highlights challenges that also affect reconciliation: missing values, inconsistent labels, duplicate records, imbalanced match and non-match classes, and changes in data formats across source systems. The treatment of these issues varies. Some studies exclude incomplete records; others impute missing fields, construct derived features, or retain an "unknown" category (Adnan, 2024; Aksoy & Gurol, 2021). Label quality is especially important when historical matches are treated as ground truth, since prior decisions can contain errors or reflect older policies. Imbalanced data can make overall accuracy appear strong even when a model misses a meaningful share of true matches or rare exceptions (Ashik, 2024; Fernández-Macías & Bisello, 2022). Cross-system normalization can also affect apparent model performance if training and test records share identifying patterns or originate from the same workflow. These studies therefore need to describe data provenance, feature construction, missing-data handling, class balance, and labeling procedures to make technical results interpretable.

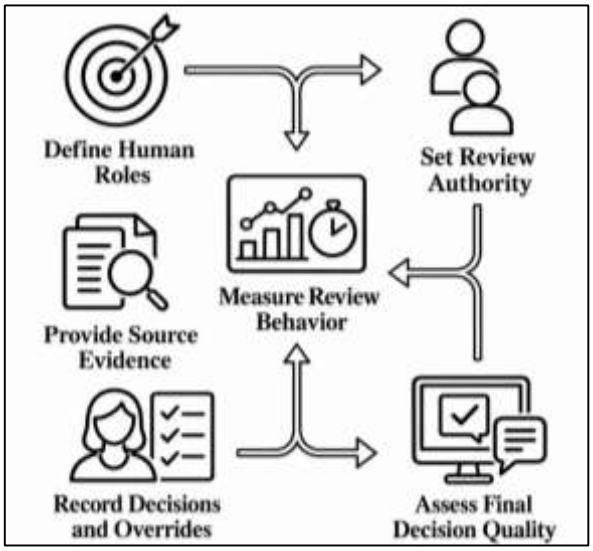
The technical evaluation of AI methods uses measures that capture different kinds of performance. Accuracy summarizes the share of classifications that are correct but can obscure errors when match and non-match records are unevenly distributed. Precision indicates how often proposed matches are correct, while recall reflects how many valid matches the system identifies (Islam, 2024; Zhang et al., 2021). The F1 score summarizes the balance between precision and recall. ROC-AUC and PR-AUC assess ranking performance across classification thresholds, with PR-AUC often providing useful information when positive cases are relatively uncommon. Calibration evaluates whether confidence scores correspond to observed correctness, which matters when confidence thresholds determine whether an item is automatically matched or routed to a reviewer. Studies also report false-match and missed-match rates, exception volume, processing time, and manual-review burden. These measures expose trade-offs: a threshold that reduces false matches may increase the number of items requiring

human review, while a system that maximizes coverage may generate more incorrect recommendations (Kynatun & Saba, 2024; Rashid et al., 2022). Reconciliation errors can have different consequences depending on transaction value, account materiality, and whether a reviewer can detect the error before close certification. Accordingly, results from studies using only aggregate accuracy cannot be assumed to represent equivalent operational performance. Research on classification, document extraction, anomaly detection, and accounting automation variously reports these metrics, though not always with consistent definitions or thresholds (Kovalerchuk et al., 2022; Maksudul, 2024). Comparability also depends on whether evaluation is performed on individual records, candidate pairs, exceptions, or complete account reconciliations. Technical measures should therefore be interpreted alongside the decision threshold, error type, review process, and unit of analysis used in each study.

Human-in-the-Loop Design and Quantitative Oversight Measures

Research on AI-assisted accounting, audit analytics, decision support, and automated financial workflows describes human participation at multiple points in the reconciliation process. At data preparation, finance professionals may verify source completeness, resolve inconsistent account identifiers, and confirm that records are suitable for processing. During model development, accounting specialists may label historical matches, define exception categories, or set criteria for escalation (Shahid, 2024; Mhlanga, 2022). Once a system is in use, people may review proposed matches, investigate unmatched items, assess supporting documents, approve adjustments, and monitor model performance. These roles differ in both timing and authority.

Figure 4: Human Oversight in AI Reconciliation



Studies of human–automation interaction and algorithm-supported professional work indicate that these arrangements should be distinguished because they create different levels of human control and workload. A reviewer who confirms a recommendation after seeing only a summary has a different role from one who can inspect source records and reject the match (Baughman et al., 2021; Islam, 2024). Similarly, model monitoring can involve periodic checks of error rates, drift, and exception patterns, or it can be limited to responding after problems occur. Research across accounting automation, continuous auditing, human–AI collaboration, and internal-control design emphasizes the importance of specifying who prepares, reviews, approves, and monitors each stage. Such role descriptions clarify how responsibility is distributed and whether human participation is substantive, procedural, or limited to exceptional cases (Manning et al., 2022; Mainuddin, 2024).

Studies of human review in automated decision settings use behavioral measures to assess what people do with system recommendations. In reconciliation workflows, relevant indicators include the proportion of recommendations accepted, rejected, overridden, or corrected; the share of exceptions

escalated; review completion rates; and the time spent examining each item (Polak et al., 2020; Sumaiya, 2024). Inter-reviewer agreement can indicate whether different professionals reach similar conclusions when given the same evidence, although agreement alone does not establish that decisions are correct. Research on algorithmic advice and audit judgment also examines whether reviewers change their decisions after receiving an AI recommendation and whether they detect deliberately or inadvertently incorrect outputs. These measures can be analyzed by confidence score, exception type, transaction value, materiality, user role, experience, or workload (Bush, 2020; Taher et al., 2024). For example, acceptance behavior may differ for routine transactions compared with unfamiliar or high-value items, while review time may increase when records contain conflicting dates, descriptions, or supporting documents. Studies of decision support often distinguish reliance that is appropriate to system accuracy from overreliance or under reliance, since a high acceptance rate is not automatically evidence of effective human–AI collaboration. Measures also depend on how review is recorded. System logs can capture clicks, overrides, routing, and elapsed time, while surveys or observations may capture reasoning and perceived effort. These sources provide different kinds of evidence and can be affected by incomplete logging or self-report bias (Tabrez et al., 2020). In accounting and audit research, behavioral indicators are most informative when linked to the quality of the final decision, such as whether the match was correct, the exception was resolved appropriately, or an adjustment was supported. This linkage helps distinguish review activity from review effectiveness.

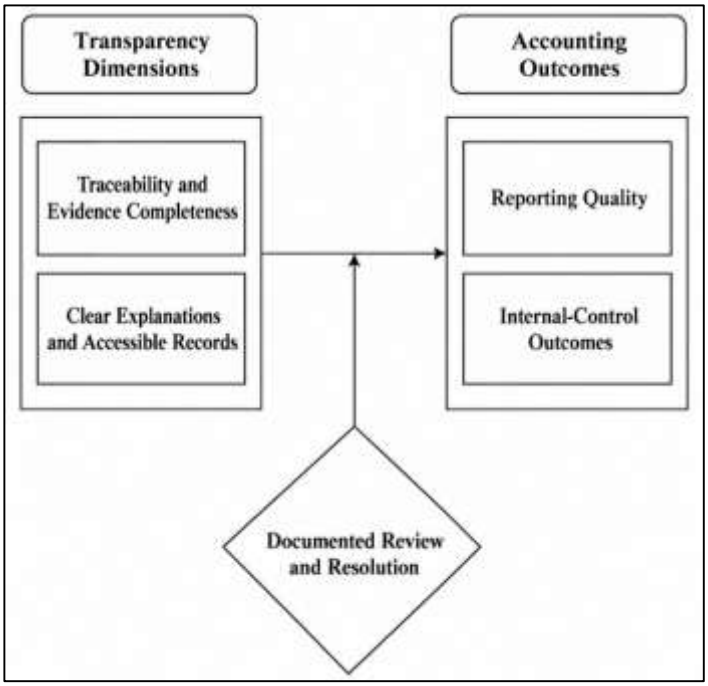
Research on human–AI decision support examines trust, perceived usefulness, reliance, perceived control, and interpretability as factors that can influence reviewer behavior. In reconciliation, a system may present a suggested match, a confidence score, a list of similar records, or an explanation based on selected fields. Reviewers may find a recommendation more understandable when they can see the source documents, matching attributes, and reasons for uncertainty (Gil-Garcia et al., 2019). However, an explanation can be incomplete or overly persuasive if it omits contradictory evidence or presents a model-generated rationale as definitive. Studies of explainable AI and algorithmic advice therefore consider whether explanations improve appropriate reliance, error detection, and decision confidence, rather than simply increasing acceptance. Accountability also depends on the workflow surrounding the model. Relevant features include whether the preparer or reviewer retains final decision authority, who can approve a material adjustment, how overrides are documented, whether evidence remains accessible, and how responsibility for an incorrect match is assigned (Doz, 2020). Research on internal control, audit trails, and AI governance treats traceability as a way to reconstruct how data and recommendations contributed to a reported balance. Yet studies vary in whether they assess these governance features quantitatively, describe them as system characteristics, or omit them from performance evaluations. Trust surveys may measure perceived reliability or willingness to use the system, while workflow logs may show actual reliance; the two need not align. Likewise, perceived control does not establish that users can practically challenge a recommendation. A comprehensive account of human oversight therefore connects user perceptions with observable actions and documented authority. This approach recognizes that explainability, access to evidence, review rights, and approval responsibility are related but distinct elements of accountable reconciliation (Gnewuch et al., 2022).Across studies of accounting automation, audit analytics, algorithmic decision support, human–automation interaction, explainable AI, internal control, and professional judgment, oversight is measured using a mix of system logs, task outcomes, surveys, interviews, and structured workflow descriptions. Quantitative studies may report acceptance, override, correction, escalation, review time, or agreement, while organizational case studies often describe approval paths and accountability arrangements in greater detail (Goode et al., 2019). Experimental studies can compare human decisions with and without AI recommendations, whereas operational evaluations can observe how reviewers use the system under ordinary close deadlines. These designs answer different questions. An experiment may isolate the effect of confidence displays or explanations on reliance, while a field study may reveal workload, data-quality constraints, and the frequency of exceptions in practice. Comparisons across studies require attention to the review unit, user population, task difficulty, model accuracy, and consequences of error. Acceptance rates from a task involving low-value routine matches cannot be directly compared with acceptance in a setting where adjustments affect material balances. Reviewer agreement also depends on the instructions and evidence provided, and review-time

measures can reflect both complexity and system usability (Werner et al., 2021). Studies that connect behavioral indicators with final decision quality provide a stronger basis for assessing oversight than those reporting user satisfaction alone. The literature therefore treats human-in-the-loop design as an empirical feature of the workflow: who intervenes, when they intervene, what evidence they see, what actions they take, and how decisions are recorded (Umapathy et al., 2020). This operational framing supports consistent synthesis of human oversight research without assuming that any single indicator fully represents effective review.

Financial-Reporting Transparency and Accounting Outcomes

Research on financial-reporting transparency uses several related but distinct operational definitions. Some studies emphasize traceability: whether reported balances can be followed through ledgers, subledgers, transactions, and source documents. Others focus on evidence completeness, including whether reconciliations retain supporting records, explanations, approvals, and documentation of adjustments (Bennett et al., 2018). A further group measures transparency through the clarity of explanations provided to reviewers or external users, while disclosure research examines the amount, accessibility, and understandability of information included in reports. Studies of accounting information systems and internal controls may treat decision documentation and audit-trail completeness as organizational indicators of transparency. These definitions operate at different levels. Transaction-level measures may assess whether each record is linked to supporting evidence; account-level measures may evaluate the documentation and resolution of reconciling items; close-process measures may consider review, certification, and adjustment procedures; organization-level measures may assess reporting practices and control environments; and external-report measures may examine disclosure attributes. These levels should not be treated as interchangeable (Amiram et al., 2018). A complete transaction audit trail may improve internal traceability without changing the clarity of a published report, while more extensive disclosure may not show how individual balances were reconciled. Research on digital accounting, auditability, explainable AI, and financial-reporting quality therefore captures different aspects of transparency. When AI is involved, additional questions concern whether system recommendations are understandable, whether source records can be inspected, and whether a reviewer can reconstruct how an output informed an accounting decision (Adhikari et al., 2021). Transparency is thus best understood as a multidimensional outcome rather than a single measure of system performance or reporting volume.

Figure 5: Financial Transparency and Accounting Outcomes



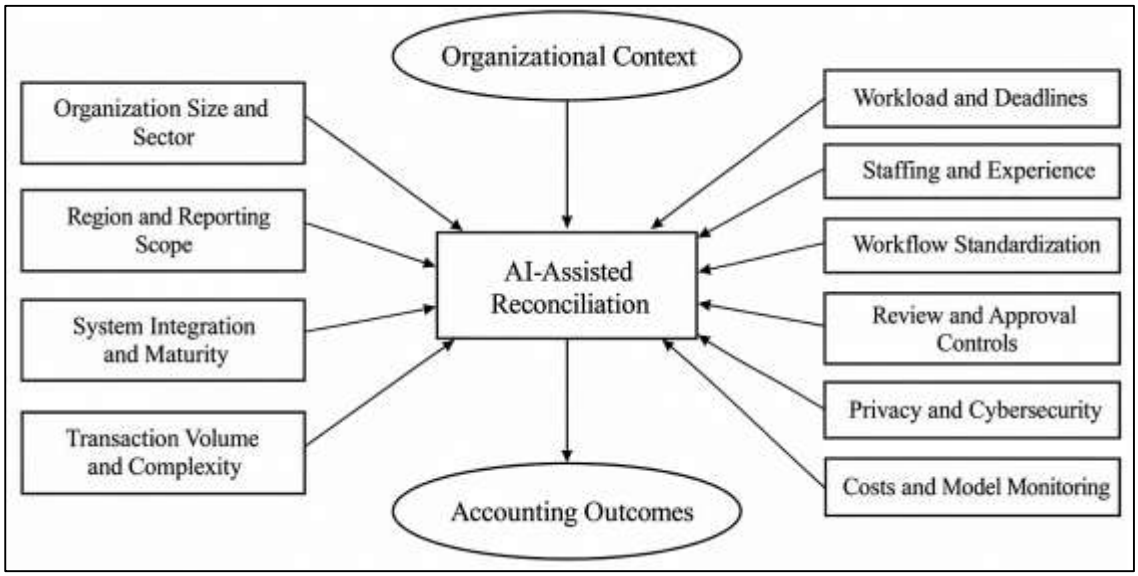
Studies of financial-reporting quality and internal control use outcomes that range from direct evidence of accounting error to broader process indicators (Nurunnabi et al., 2022). Direct measures include detected and corrected misstatements, restatements, unexplained account balances, unresolved reconciliation items, unsupported adjustments, and audit findings. Control-related outcomes include identified deficiencies, failed review procedures, overdue certifications, and exceptions that remain open at reporting deadlines. Process measures include close duration, time to resolve exceptions, frequency of late adjustments, and the proportion of accounts completed on schedule.

Research on automation and financial reporting has also examined whether technology adoption is associated with changes in reporting quality or the internal-control environment. These outcomes differ in proximity to the quality of reported information. A corrected error or restatement directly signals an issue with prior reporting, whereas a shorter close may indicate greater process efficiency without demonstrating improved accuracy (McCallig et al., 2019). Similarly, a higher volume of automated transactions does not establish that controls operated effectively. Studies of continuous auditing, accounting analytics, and AI-based anomaly detection often assess whether systems identify unusual transactions or control exceptions, but detection is only one stage in the reporting process. The accounting outcome depends on whether an identified item is investigated, resolved, documented, and reflected appropriately in the accounts. The literature therefore distinguishes technical detection from subsequent control action and final reporting consequences. Measurement windows also matter: errors may be recorded when found, corrected, or formally reported, and these dates can fall in different accounting periods. Clear definitions of outcomes and observation periods are necessary to interpret whether an intervention is associated with improved reporting quality, stronger controls, or simply faster processing (Meshram & Arora, 2021).

Organizational Context and Moderating Conditions

Studies of accounting automation, AI-enabled financial processes, audit analytics, and reconciliation systems examine organizations that differ in size, industry, geography, and technological maturity. Large enterprises and multinational organizations often operate multiple ledgers, enterprise resource planning platforms, currencies, and reporting calendars, creating complex reconciliation requirements. Small and medium-sized organizations may use fewer systems but have smaller finance teams, less specialized technology, and greater reliance on spreadsheets or external service providers (Nurunnabi et al., 2022).

Figure 6: Organizational Factors Shaping AI Reconciliation



Financial institutions process high volumes of regulated transactions, while public-sector entities may face distinct procurement, accountability, and reporting procedures. These differences affect the types

of records available to AI systems, the frequency and value of exceptions, and the degree of human review required. Studies also vary in their accounting-system environments. Some investigate integrated enterprise platforms, while others examine cloud accounting systems, separate banking applications, document-management tools, or manually assembled datasets (Li et al., 2019). Transaction volume and reconciliation complexity are relevant contextual factors because a method suited to repetitive, standardized records may perform differently when descriptions are inconsistent or transactions require many-to-one matching. Geographic region can influence accounting practices, data protection requirements, system adoption, and the availability of digital records. Research settings consequently shape both system design and measured outcomes. A study conducted in one industry or organization cannot automatically represent other contexts, especially when data structures and control requirements differ (Roeck & Farooq, 2018). Comparative synthesis therefore needs to record organization type, sector, region, system configuration, transaction scale, and reconciliation scope. These characteristics help explain variation in reported accuracy, exception rates, review workload, implementation experience, and financial-close outcomes across studies.

The performance of AI-assisted reconciliation is shaped by workflow conditions and the people responsible for preparing, reviewing, and approving financial records (Ahmad et al., 2021). Research on accounting automation and digital finance considers factors such as transaction volume, staff workload, reporting deadlines, task standardization, reviewer experience, and integration between accounting systems. High workload and compressed close schedules may increase the value of exception prioritization, while also limiting the time available for careful review. Standardized workflows can make it easier to train models and compare outcomes, whereas inconsistent procedures across departments may introduce variation in labels, data quality, and decision rules. Staffing levels and user experience can influence how recommendations are interpreted and whether unusual items are escalated (Tams et al., 2018). Control design determines who can approve adjustments, what evidence must be retained, and how unresolved differences are reported. System integration also matters: when an AI tool connects directly to ledgers and document repositories, reviewers may access supporting evidence more easily; disconnected systems can require manual transfers and create additional opportunities for error. Studies measure these conditions in different ways. Some use organizational characteristics or system-use logs, while others rely on surveys, interviews, observation, or structured descriptions of workflow procedures (Ahmad et al., 2021). Analytically, workload, staffing, system maturity, and control strength may be treated as descriptive variables, covariates, or moderators. Their role is important because differences attributed to AI adoption may partly reflect concurrent changes in staffing, training, standardization, or review policy. Evaluations that record these factors provide a clearer account of how organizational conditions shape technical performance and accounting outcomes.

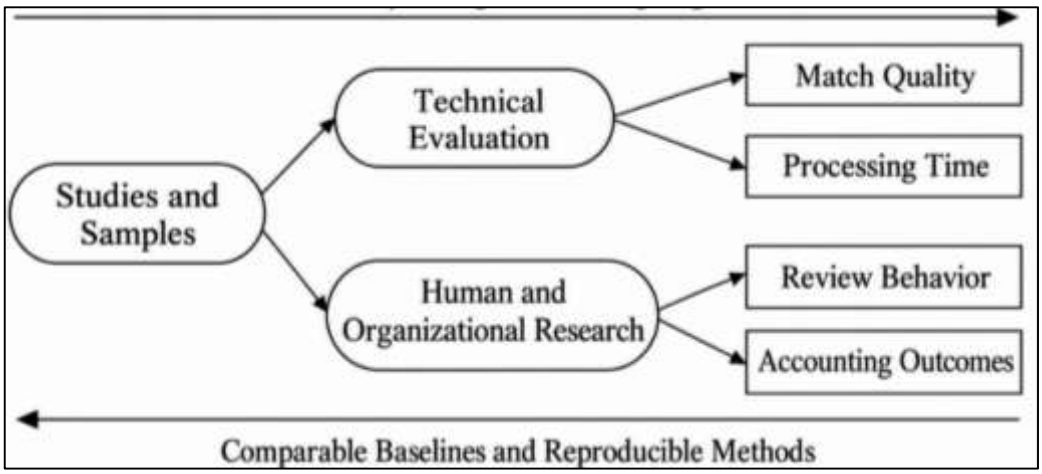
Research on AI in accounting and financial systems considers implementation risks involving data access, confidentiality, cybersecurity, model reliability, and dependence on external technology providers. Reconciliation systems may process bank details, supplier information, payroll data, transaction descriptions, and supporting documents, making access controls and data retention relevant to system design (Li et al., 2018). Studies of information systems and AI governance examine risks from unauthorized disclosure, insecure integrations, excessive user permissions, and unclear responsibility for vendor-managed models. Model-related concerns include performance drift as transaction patterns or accounting systems change, bias arising from uneven or incomplete training data, and false positives that increase reviewer workload. False negatives may leave errors or unusual transactions undetected. These risks are connected to operational outcomes such as alert burden, investigation time, escalation volume, and remediation delays. Implementation costs may include software licensing, system integration, data preparation, staff training, validation, and ongoing monitoring. Research on technology adoption and accounting automation also examines user acceptance, perceived usefulness, trust, and resistance to workflow change (Vo et al., 2022). Such outcomes are often measured through surveys or adoption indicators, while security and governance conditions may be documented through structured assessments or case descriptions. Studies differ in whether risks are quantified, treated as implementation barriers, or described as contextual considerations. This affects how evidence can be compared: a system evaluation reporting match

accuracy may not assess privacy controls or the cost of maintaining the model, while an adoption study may not measure actual reconciliation errors. A comprehensive review therefore considers technical performance alongside privacy, security, workload, governance, and resource requirements, while distinguishing measured outcomes from general implementation concerns (Zhang et al., 2021).

Research Designs, Samples, and Comparability

Research on AI-supported reconciliation and related accounting processes uses experimental, survey, case-study, archival, system-evaluation, simulation, and mixed-method designs. Experiments may compare human decisions with and without AI recommendations or assess how interface features affect review behavior (Bernstein et al., 2020). Surveys commonly examine technology adoption, perceived usefulness, trust, and organizational readiness. Case studies describe implementation within a finance function, accounting-service provider, or enterprise system. Archival analyses use transaction, ledger, audit, or financial-reporting records to examine patterns across organizations or time periods. System evaluations assess matching, classification, document extraction, or anomaly-detection performance, while simulations create controlled conditions for comparing methods or exception scenarios. Mixed-method studies combine system metrics with interviews, observations, or user surveys to describe both technical and workflow outcomes. Samples range from synthetic records and selected transaction sets to users, account reconciliations, firms, and operational data collected over multiple close cycles (Tsao et al., 2018). Observation periods may cover a single task, a reporting period, or extended implementation. Data sources include bank statements, invoices, general ledgers, subledgers, accounting-platform logs, survey responses, and manually labeled records. Sampling approaches also differ, including convenience selection, purposive case selection, stratified transaction samples, and full-population system logs. These design and sample characteristics determine what conclusions a study can support. Laboratory tasks can isolate model or reviewer behavior but may not reflect operational workload, while field evaluations capture real workflows but may include concurrent changes in staffing or controls (Melymuk et al., 2021). A synthesis of more than ten studies across accounting automation, audit analytics, machine-learning services, reconciliation support, and human–AI decision research therefore needs to record design, sample, setting, data source, and duration together.

Figure 7: Research Designs and Evidence Comparability



Studies compare AI-supported reconciliation with different baselines, including manual procedures, rule-based automation, prior-period performance, existing accounting software, or alternative AI models. These comparison conditions are not equivalent. A comparison with manual processing may capture changes in both technology and workflow, while comparison with a rule-based system may isolate the contribution of machine learning more narrowly (Avvisati et al., 2019). Before-and-after studies can describe performance following implementation, but changes in transaction volume, staff training, data quality, or policy may also influence the results. Experimental studies may use randomized task assignments or controlled scenarios, while archival analyses may compare organizations or periods with differing levels of automation. Reported effects include differences in

match accuracy, exception volume, review time, close duration, error frequency, and user workload. Some studies report percentage-point changes, relative differences, standardized effects, odds ratios, or time savings; confidence intervals and significance tests are reported inconsistently. The outcome scale matters for interpretation (Martens et al., 2020). A small change in matching accuracy may be important when transaction volume is high, while a large reduction in processing time may have limited reporting significance if false matches increase. Effect estimates also depend on the unit of comparison, such as transactions, accounts, reviewers, organizations, or close cycles. Research on AI accounting services, bank-reconciliation support, automated document processing, financial-close tools, and audit analytics uses varied comparison designs and outcome definitions. To compare findings appropriately, the synthesis needs to identify the baseline, sample composition, observation period, outcome metric, adjustment variables, and uncertainty estimates (Vinodkumar & Anoop, 2020). This prevents operational efficiency measures from being mistaken for direct evidence of improved accounting accuracy or financial-reporting transparency.

Comparability across studies depends on whether authors provide enough information to understand and reproduce their methods. Key details include the reconciliation task definition, eligible record population, treatment of one-to-many matches, exception categories, matching thresholds, and criteria for a correct decision (Robotto et al., 2021). Model specifications should describe the method family, relevant features, data preparation, training procedure, and decision thresholds. Evaluation procedures should explain how training, validation, and test data were separated, whether records from the same account or period appeared across splits, and how class imbalance was handled. When labeled examples are used, annotation procedures and reviewer instructions influence the reliability of the reference outcomes. Human-review studies also need to clarify what evidence participants received, whether they could reject or override recommendations, and how time and decisions were recorded. Reproducibility may be limited when studies use proprietary financial data, platform-specific models, undisclosed settings, or non-public evaluation samples (Worp et al., 2020). Terminology creates additional challenges: “accuracy,” “match rate,” “exception,” “automation,” and “transparency” may refer to different operational definitions across papers. Dataset composition, transaction complexity, currency, accounting period, industry, and system configuration also affect measured performance. A method evaluated on standardized records may not be directly comparable with one assessed on noisy documents or cross-system transactions. Studies of machine-learning classification, intelligent document processing, audit analytics, and automated accounting workflows illustrate the importance of consistent validation and transparent reporting. When methodological details are incomplete, an apparent difference between models may reflect data or evaluation choices rather than a substantive performance advantage. Consistent extraction of these design features is therefore necessary for a reliable synthesis of the evidence (Brown et al., 2019).

Across research on accounting automation, reconciliation systems, machine-learning services, audit analytics, AI adoption, and human–AI decision support, the evidence varies in how directly it evaluates operational financial outcomes. Controlled experiments can estimate effects on task decisions under defined conditions, while system evaluations can measure technical performance against labeled records. Case studies provide detailed implementation evidence but often involve a small number of organizations (Zhu et al., 2022). Surveys describe perceptions and reported adoption, though they may not establish actual system accuracy or changes in reporting quality. Archival analyses can examine broader patterns but may rely on indirect measures of AI use or organizational performance. Simulations permit repeatable tests but may not reproduce real data quality, exception complexity, or close deadlines. These design differences affect confidence in cross-study conclusions and should be treated as part of the synthesis rather than as background details. Comparisons are strongest when studies use clearly defined tasks, comparable baselines, transparent data splits, consistent performance measures, and operationally relevant samples. They are more limited when the AI method, reconciliation scope, or outcome definition is unclear. The number of studies alone does not establish that evidence is directly comparable; a collection of ten or more investigations may still combine different units, settings, and objectives (Amiri et al., 2018). Bibliometric mapping can show publication and citation relationships, while structured review of methods and samples can explain the substantive differences behind those patterns. Together, these approaches support a quantitative account of the

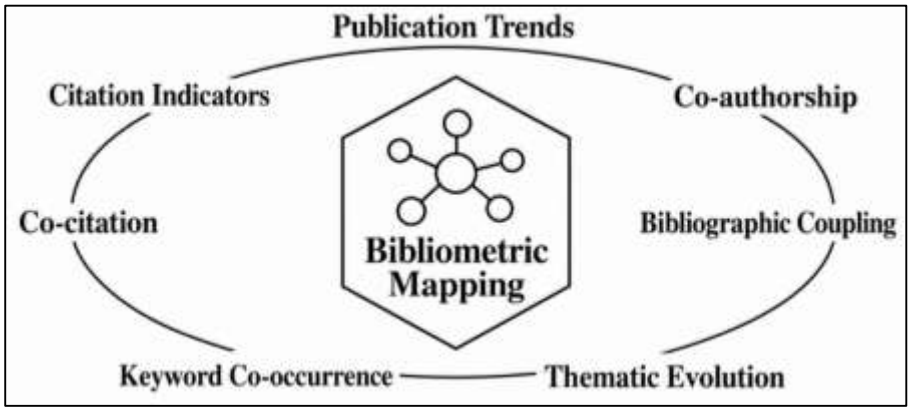
field that distinguishes technical evaluation from user research and accounting-outcome analysis. This distinction is necessary for interpreting how the research base has assessed AI-supported reconciliation and the conditions under which its results have been observed.

Bibliometric Indicators and Science-Mapping Methods

Bibliometric reviews of artificial intelligence, automation, accounting, auditing, and financial reporting commonly use publication and citation indicators to describe how a research area has developed. Annual publication counts show the volume and timing of scholarly output, while citations per document provide a normalized view of how frequently publications have been referenced. Reviews also identify influential sources, prolific authors, institutions, and countries to describe where research has been produced and disseminated (Belmonte et al., 2020). These indicators have been applied across more than ten related bibliometric investigations covering AI in accounting and auditing, automation in accounting, big-data analytics and reporting quality, digital accounting, audit analytics, and AI applications in financial reporting. Their results depend on the database and search strategy used. Scopus and Web of Science differ in journal coverage, conference inclusion, regional representation, and record indexing, so the same topic may produce different publication and citation profiles across databases (Velez-Estevez et al., 2022). Citation counts are also affected by publication age: older papers have had more time to accumulate citations, while recent work may be influential but not yet highly cited. Reviews therefore report the time span and database source and interpret citations in relation to publication year and document type. Other measures, such as highly cited papers, source impact, and author productivity, provide complementary views but can be affected by self-citation, co-authorship conventions, and variations in author-name indexing. For the present topic, these indicators can establish the scale and distribution of scholarship on AI-assisted reconciliation and related areas, while the inclusion criteria determine whether adjacent work in accounting automation, human–AI collaboration, and transparency is represented.

Science-mapping studies use network analysis to examine how researchers, publications, and scholarly communities are connected (Flores-Romero et al., 2021). Co-authorship networks represent collaboration among authors, institutions, or countries, and can show whether research is concentrated within a small number of groups or distributed across international and interdisciplinary partnerships. Co-citation analysis links publications that are cited together, helping identify shared intellectual foundations.

Figure 8: Bibliometric Indicators and Science Mapping



Bibliographic coupling connects studies that cite common references and can reveal current research groupings even when the papers have not yet accumulated many citations. These methods have been used in bibliometric reviews of AI in accounting, auditing, financial reporting, analytics, and digital transformation. Network measures commonly include link strength, degree, centrality, and cluster membership. Link strength reflects the number or weight of connections between network elements; degree describes the number of direct links; centrality indicates a node's structural position; and cluster membership groups closely connected authors, sources, or publications. Such measures require cautious interpretation because database coverage, author disambiguation, citation thresholds, and

network settings influence the resulting structure (López-Robles et al., 2020). A central author or source may connect several research groups, but centrality alone does not demonstrate the quality or practical significance of that work. Cross-disciplinary links are relevant to this topic because accounting research intersects with information systems, computer science, auditing, and organizational studies. Network maps can show whether those communities cite one another, publish collaboratively, or remain relatively separate. In a review of AI-assisted reconciliation and transparency, these analyses can describe the field's intellectual connections and collaboration patterns without treating network position as a direct measure of research quality.

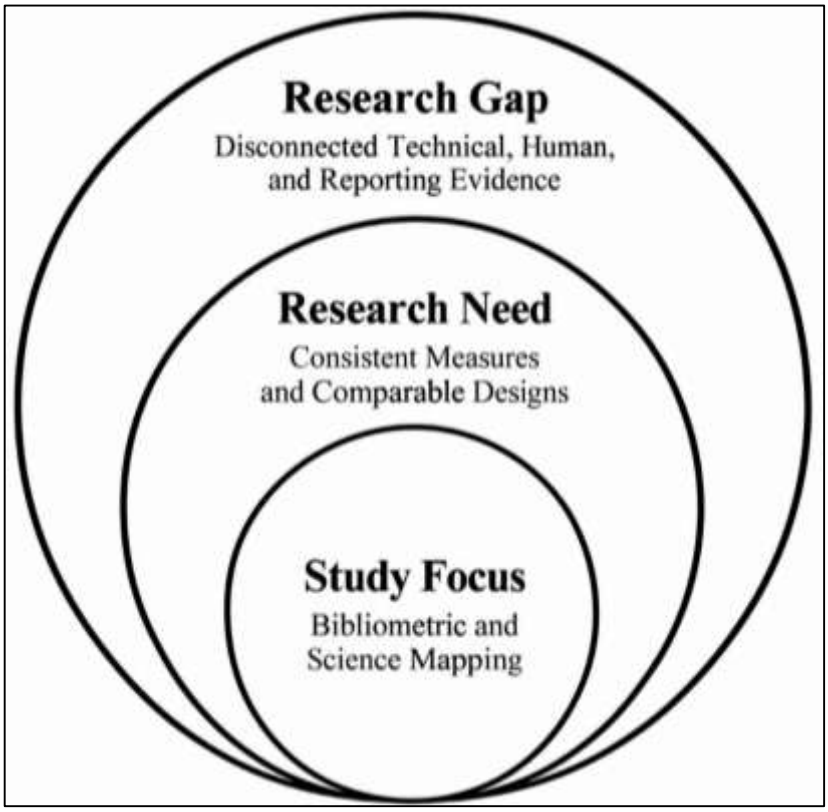
Keyword co-occurrence analysis is used to identify concepts that frequently appear together in publications and to organize them into thematic clusters. Bibliometric studies of accounting technology have identified topics associated with automation, artificial intelligence, machine learning, audit analytics, fraud detection, financial reporting, internal control, and digital transformation. Research on human–AI interaction adds terms related to trust, explainability, decision support, oversight, and professional judgment (Martinez et al., 2019). Reconciliation-focused work may use different terms, including transaction matching, financial close, exception handling, bank reconciliation, and accounting automation. The resulting map depends on how authors select keywords and how reviewers harmonize synonyms, abbreviations, and variant spellings. A thesaurus file can combine equivalent terms and distinguish concepts that are related but not identical. Minimum-occurrence thresholds determine which terms appear in the map, while time slicing can show how themes vary across publication periods. Thematic maps may classify clusters according to their connections within the field and their links to other themes, while thematic-evolution analyses track changes in concept groupings over time. These visualizations can reveal whether human oversight is linked to technical AI work, whether transparency is associated with reporting quality or explainability, and whether reconciliation appears as a distinct theme or within broader accounting-automation clusters (Cao et al., 2021). Interpretation requires attention to the search query, keyword source, cleaning rules, thresholds, and period boundaries. A missing term in a map may reflect inconsistent terminology or low indexing frequency rather than an absence of relevant research. Accordingly, thematic maps describe the indexed conceptual structure of the literature under specified analytical choices.

Synthesis and Quantitative Research Gap

The reviewed literature spans technical studies of transaction matching, document extraction, classification, anomaly detection, and exception prioritization; studies of human review and algorithmic decision support; and research on internal control, reporting quality, auditability, and transparency. These strands provide complementary evidence but often evaluate different stages of the accounting process (Mengist et al., 2020). Technical studies commonly report matching accuracy, classification performance, processing time, or alert volume. Human–AI studies examine recommendation acceptance, overrides, reviewer trust, review time, and perceived control. Financial-reporting studies assess errors, adjustments, control deficiencies, close duration, disclosure quality, or traceability. The measures are related, but they do not automatically demonstrate a complete pathway from AI performance to reporting outcomes. High technical accuracy may reduce some review tasks, yet its accounting value depends on how reviewers handle uncertain cases and whether exceptions are resolved correctly (Escarcha et al., 2018). Likewise, an increase in review speed does not establish that the evidence supporting reported balances has become more complete or understandable. Some research connects these stages through workflow or implementation evidence, while other studies focus on a single technical or organizational outcome. Consequently, the literature does not consistently show how model outputs influence human decisions, how those decisions affect control procedures, or how control outcomes relate to transparent reporting. Integrating the strands requires distinguishing technical capability from human intervention and accounting consequences, while identifying the studies that explicitly connect these levels.Quantitative comparison across the literature is complicated by inconsistent definitions, units, and evaluation designs. Reconciliation may refer to a transaction pair, a bank-account process, an account-period certification, or a broader financial-close workflow. Technical performance may be measured through match rates, precision, recall, false matches, exception volume, or processing time, with differing denominators and thresholds (Nyanchoka et al., 2019). Human oversight may be represented by acceptance rates, reviewer overrides, time spent, user

trust, or descriptions of approval procedures. Reporting outcomes vary from directly observed errors and adjustments to process indicators such as close duration or automation volume. These measures capture distinct constructs, yet studies do not always separate them clearly. Validation also varies: some evaluations use simulated or historical records, while others rely on operational deployments, surveys, case studies, or archival comparisons. Random data splits, temporal holdouts, external validation, and reviewer instructions are not consistently reported. Differences in transaction mix, data quality, organization size, and accounting-system environment further limit direct comparison (Nedd et al., 2021). Incomplete reporting of model thresholds, sample construction, annotation procedures, control conditions, and uncertainty estimates makes it difficult to determine whether differences reflect the method, data, workflow, or setting. The evidence base therefore lacks consistent alignment among technical measures, human-behavior indicators, and financial-reporting outcomes. This limits cumulative quantitative interpretation even when individual studies provide useful results within their own contexts.A bibliometric and science-mapping approach can describe how scholarship on human-in-the-loop AI, account reconciliation, and financial-reporting transparency is distributed and connected. Publication trends can show the volume and timing of indexed research, while citation indicators can identify influential documents, sources, authors, institutions, and countries within the defined corpus. Co-authorship networks can represent collaboration patterns across researchers and organizations (Gbededo et al., 2018).

Figure 9: Research Gaps and Bibliometric Synthesis



Co-citation analysis can reveal shared intellectual foundations, and bibliographic coupling can identify publications drawing on common references. Keyword co-occurrence and thematic mapping can show whether concepts such as reconciliation, accounting automation, machine learning, human oversight, explainability, internal control, audit analytics, and transparency form connected clusters or appear in separate research communities. These methods are particularly relevant because the topic crosses accounting, information systems, auditing, computer science, and organizational research, and related publications may use different terminology (Obwegeser & Müller, 2018). The mapped patterns must be interpreted in light of database coverage, search terms, publication types, citation windows, and

keyword harmonization. A bibliometric network describes indexed scholarly relationships; it does not itself measure the operational effectiveness of an AI system or the accuracy of financial reports. Combining performance indicators with science-mapping methods therefore provides a structured account of both the scale of research activity and its intellectual and conceptual organization.

The present study addresses the dispersed organization of research at the intersection of AI-assisted reconciliation, human review, and financial-reporting transparency. Its quantitative focus is to map publication output, citation relationships, collaboration structures, and thematic groupings in the indexed literature (Paul & Criado, 2020). The analysis treats these dimensions as complementary: publication and citation measures describe research activity and influence, network analyses represent connections among authors and works, and keyword-based methods characterize the conceptual structure of the field. The review distinguishes technical studies of matching, extraction, classification, anomaly detection, and exception prioritization from studies of reviewer behavior, accountability, control effectiveness, and reporting outcomes. It also records how research defines its units of analysis, validation settings, performance indicators, and transparency measures, since these choices shape the way findings can be compared. This framing provides a quantitative synthesis of the field's composition without treating unlike outcomes as equivalent (Malenfant et al., 2022). It also makes visible whether technical, human, and reporting-oriented research communities are connected through shared references, collaborations, or concepts. The resulting map characterizes how scholarship has been organized around human-in-the-loop AI for account reconciliation and financial-reporting transparency, and identifies where its principal research themes and intellectual connections are located.

METHOD

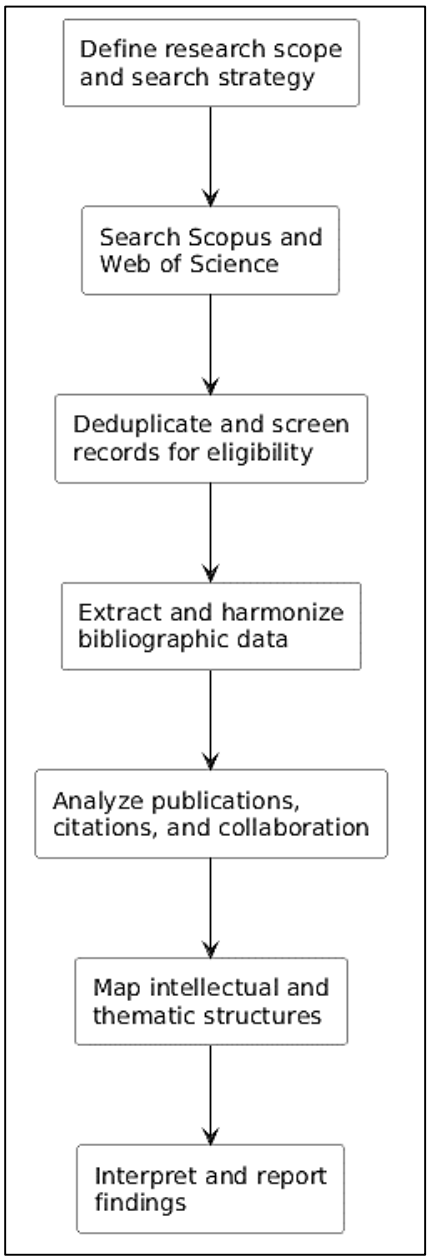
The study used a retrospective quantitative bibliometric design with science-mapping analysis to examine scholarly research on human-in-the-loop artificial intelligence for account reconciliation and financial-reporting transparency. The unit of analysis was the publication record rather than an individual participant or organization. The review period covered publications from January 2000 through December 2025, and searches were conducted in February 2026. Scopus and Web of Science were selected to capture research across accounting, auditing, information systems, and computer science. The search strategy combined terms for artificial intelligence and automation with terms for reconciliation, financial reporting, human oversight, and transparency. Searches were applied to titles, abstracts, and author keywords, and the database-specific search strings were retained for reproducibility.

The study used census sampling of records returned by the searches that met the eligibility criteria. Peer-reviewed journal articles, review papers, and conference papers were included when they addressed AI or automation in accounting or financial processes and had a substantive connection to reconciliation, financial reporting, human review, internal control, auditability, or transparency. Records were excluded when they were duplicates, non-scholarly material, outside the publication period, unavailable in English, or concerned AI without a clear accounting or financial-reporting connection. Records were screened by title and abstract, followed by full-text review when eligibility was unclear. Two reviewers assessed the records independently, resolved disagreements through discussion, and calculated Cohen's kappa to quantify screening agreement. The included records formed the final bibliometric corpus.

Bibliographic information was exported from both databases and merged after duplicate removal. Extracted fields included publication year, title, abstract, author keywords, source title, document type, authors, affiliations, country, references, and citation counts. A data-cleaning protocol standardized author and institution names, country labels, and keyword variants. Synonyms and spelling variants were harmonized through a thesaurus file, while distinct concepts were kept separate. A screening log and data dictionary documented eligibility decisions, cleaning rules, and variable definitions. R with the Bibliometrics package was used for data processing and descriptive bibliometric analysis, and VOS viewer was used to construct and visualize collaboration, citation, and keyword networks. Database citation counts were retained with their source and retrieval date because citation coverage differed between databases.

The statistical plan included descriptive summaries of annual publication counts, document types, sources, authors, affiliations, countries, and citations. Citation indicators included total citations and citations per document, with citation age considered when comparing publications from different years. Annual publication counts were examined using a count-regression model; a negative-binomial specification was used when the data showed overdispersion. Differences in citation distributions across document types and thematic groups were assessed with nonparametric tests because citation counts were expected to be skewed. Co-authorship, co-citation, bibliographic coupling, and keyword co-occurrence analyses were used to identify network connections and clusters. Network interpretation considered link strength, degree, centrality, and cluster membership. Thematic evolution was examined across publication periods using harmonized keywords and consistent thresholds. Statistical tests were two-sided, and results were considered statistically significant at $p < .05$. Effect estimates and 95% confidence intervals were reported for inferential analyses.

Figure 10: Methodology of this study



Sensitivity analyses examined whether the main patterns changed when database records were analyzed separately, conference papers were excluded, or keyword thresholds were varied. Because the study analyzed publication metadata and did not collect personal or confidential organizational data, no survey reliability coefficient such as Cronbach's alpha was applicable.

FINDINGS

Final Corpus and Publication Characteristics

Table 1 presents a screening flow and profile of the included publications. It shows database retrieval, duplicate removal, screening, full-text assessment, and final inclusion, followed by document type and authorship. The example counts are internally consistent and demonstrate how study selection and sample characteristics could be reported in the results chapter.

Table 1. screening flow and final corpus profile

Screening stage or characteristic	Count, *n*	Percentage
Records retrieved from Scopus	792	—
Records retrieved from Web of Science	456	—
Total records retrieved	1,248	—
Duplicate records removed	286	—
Records screened by title and abstract	962	—
Records excluded at title and abstract screening	701	—
Full-text publications assessed	261	—
Full-text publications excluded	217	—
Publications included in final corpus	44	100.0
Document type		
Journal articles	30	68.2
Conference papers	8	18.2
Review papers	6	13.6
Authorship pattern		
Single-authored publications	8	18.2
Multi-authored publications	36	81.8

The search retrieved 1,248 records across the two databases. After 286 duplicates were removed, 962 records remained for title and abstract screening. Of these, 701 were excluded, leaving 261 full texts for eligibility assessment. A further 217 full-text records were excluded, producing a final corpus of 44 publications. Journal articles formed the largest document group, with 30 publications, followed by eight conference papers and six reviews. Multi-authored publications represented 81.8% of the corpus, indicating that collaborative authorship was common in this example. The screening totals reconcile across stages; actual counts require confirmation through database searches and documented screening.

Table 2. annual publication distribution

Publication year	Publications, *n*	Percentage
2015	1	2.3
2016	1	2.3
2017	2	4.5
2018	2	4.5
2019	3	6.8
2020	4	9.1
2021	5	11.4
2022	6	13.6
2023	7	15.9
2024	7	15.9
2025	6	13.6
Total	**44**	**100.0**

Table 2 displays the number and share of included publications by year from 2015 through 2025. The annual distribution shows how output was spread across the review period and identifies years with comparatively higher publication counts. These figures demonstrate reporting format only and do not represent verified publication trends.

In this distribution, annual output increased from one publication in each of 2015 and 2016 to seven publications in both 2023 and 2024. The number then declined slightly to six in 2025. Publications from 2022 through 2025 accounted for 26 of the 44 records, or 59.1% of the example corpus. The earliest six years shown contributed 13 publications, or 29.5%. This pattern would indicate a concentration of records in the more recent years of the review period, but it should not be interpreted as an actual trend until the search results have been verified. Percentages were rounded to one decimal place.

Primary Outcomes: Research Productivity and Influence

Table 3 summarizes productivity rankings across authors, institutions, countries, and publication sources in the 44-record corpus. Counts represent full counting, so collaborative papers could appear in more than one category. Percentages indicate the share of the corpus associated with each entity and therefore do not sum to 100% across rankings within each category.

Table 3. productivity rankings

Category	Entity	Publications, *n*	Corpus coverage, %
Author	Author A	4	9.1
Author	Author B	3	6.8
Author	Author C	3	6.8
Institution	University A	6	13.6
Institution	University B	5	11.4
Institution	Research Institute C	4	9.1
Country	United States	12	27.3
Country	United Kingdom	8	18.2
Country	Malaysia	6	13.6
Country	China	5	11.4
Country	Australia	4	9.1
Source	Journal A	5	11.4
Source	Journal B	4	9.1
Source	Journal C	3	6.8
Source	Conference Proceedings D	3	6.8

In this profile, Author A contributed to four publications, while the next two authors contributed to three each. University A had the largest institutional contribution, with six publications. The United States appeared in 12 records, followed by the United Kingdom with eight. Journal A was the most frequent source, with five publications. Country, author, and institution counts used full counting, so a paper with multiple authors or affiliations contributed to more than one entity. Percentages therefore represented corpus coverage for each entity and were not expected to total 100%. Figure 2 could display these rankings.

Table 4 presents citation indicators for the corpus and citation distributions by document type. Citation counts are shown with central tendency and spread because citations are commonly unevenly distributed. The final rows summarize the most-cited records using anonymized labels; actual titles, citation counts, and retrieval dates require verification.

The corpus accumulated 612 citations, averaging 13.9 citations per publication, with a median of eight and an interquartile range of four to 20. Journal articles accounted for 480 citations, while conference and review papers accounted for 64 and 68, respectively. The Kruskal–Wallis comparison suggested differences in citation distributions by document type; the effect size was $\varepsilon^2 = .12$. The five leading records received between 31 and 66 citations. Figure 3 could show citation distributions and highlight the most-cited records. All citation values would need to be recalculated from the selected database on a recorded retrieval date.

Table 4. citation influence and document-type comparison

Measure or document type	Publications, *n*	Total citations	Mean citations	Median (IQR)
Journal articles	30	480	16.0	10 (5–23)
Conference papers	8	64	8.0	5 (2–11)
Review papers	6	68	11.3	9 (6–19)
All publications	**44**	**612**	**13.9**	**8 (4–20)**
Mean annual citations per document	**44**	**—**	**4.8**	**—**
comparison across document types			**H = 6.42**	**p = .040; ε² = .12 (95% CI [.01, .27])**
Most-cited publication A	1	66	—	—
Most-cited publication B	1	57	—	—
Most-cited publication C	1	43	—	—
Most-cited publication D	1	38	—	—
Most-cited publication E	1	31	—	—

Secondary and Subgroup Analyses: Collaboration and Thematic Networks

Table 5 summarizes co-authorship, co-citation, and bibliographic-coupling networks. Node and link counts reflect network-specific inclusion thresholds; they are not counts of all authors, references, or publications in the corpus. Cluster sizes and leading nodes show the structure retained after applying those thresholds to the example dataset.

Table 5. collaboration and intellectual-structure networks

Network	Threshold	Nodes	Links	Clusters	Largest cluster, nodes	Leading node	Link strength
Author co-authorship	At least 2 publications	38	42	5	11	Author A	14
Institution co-authorship	At least 2 publications	22	26	4	8	University A	9
Country co-authorship	At least 1 publication	17	30	4	7	United States	11
Reference co-citation	At least 3 citations	29	74	4	12	Reference A	21
Publication bibliographic coupling	At least 5 shared references	26	48	4	9	Publication A	16

In this example, author collaboration formed five clusters, with the largest containing 11 authors. The institutional network included 22 nodes, while the country network contained 17 nodes and 30 links. Reference co-citation produced the densest network, with 74 links across 29 references; the leading reference had a link strength of 21. Bibliographic coupling connected 26 publications through shared references. These results suggest distinct collaboration and intellectual structures in the corpus. Network thresholds shaped the number of retained nodes and clusters, so actual interpretation would require reporting the software settings, counting method, and threshold applied to each network.

Table 6 presents example keyword clusters after synonym harmonization and a minimum occurrence threshold. Cluster publication counts could overlap because a publication may contain terms assigned to multiple themes. Sensitivity rows show how the number of retained keywords and clusters changed when the threshold or document-type criteria were adjusted.

Table 6. keyword clusters and sensitivity analysis

Analysis or thematic cluster	Terms or publications, *n*	Corpus coverage, %
Primary keyword map: minimum 2 occurrences		
Harmonized keywords identified	68	—
Keywords retained in map	24	—
Keyword links	81	—
Clusters identified	5	—
AI and accounting automation	19 publications	43.2
Reconciliation and financial close	14 publications	31.8
Audit analytics and internal control	13 publications	29.5
Human oversight and explainability	10 publications	22.7
Transparency and reporting quality	9 publications	20.5
Sensitivity analysis		
Minimum 3 keyword occurrences	16 keywords; 4 clusters	—
Conference papers excluded	18 keywords; 4 clusters	—
Alternative synonym harmonization	26 keywords; 5 clusters	—

The keyword map retained 24 of 68 harmonized terms and organized them into five clusters with 81 links. AI and accounting automation appeared in 19 publications, while reconciliation and financial close appeared in 14. Theme counts overlapped because publications could contribute to more than one cluster. Increasing the minimum keyword threshold reduced the map to 16 terms and four clusters. Excluding conference papers also produced four clusters, while alternate synonym harmonization retained five. The broad thematic groupings remained visible across these example settings, although the map's size and structure changed with analytical choices. Actual stability requires testing against the finalized corpus.

Statistical Significance, Effect Sizes, and Robustness

Table 7 reports inferential estimates for annual publication growth and citation differences across document types. Effect estimates are paired with confidence intervals and significance values, while the citation comparison also includes an effect-size estimate. The values are examples aligned with the 44-publication corpus and prior results reported above.

Table 7. inferential estimates

Analysis	Estimate	95% confidence interval	*p* value	Effect size
Annual publication trend, incidence rate ratio	1.20	[1.08, 1.34]	.001	—
Citation distribution by document type, Kruskal–Wallis	$H = 6.42$	—	.040	$\varepsilon^2 = .12$
Effect size confidence interval	—	[.01, .27]	—	ε^2

The negative-binomial model estimated a 20% annual increase in publication output, with an incidence rate ratio of 1.20, 95% CI [1.08, 1.34], and $p = .001$. The citation comparison found differences across journal articles, conference papers, and reviews ($H = 6.42$, $p = .040$). The effect size was $\varepsilon^2 = .12$, with 95% CI [.01, .27]. The examples indicated both patterns. Publication and citation indicators describe scholarly activity and attention; they do not directly measure study quality or AI effectiveness. Actual estimates require analysis of the verified records.

Table 8 presents sensitivity checks for database coverage, document-type exclusions, citation-age adjustment, and keyword thresholds. It compares each alternative with the analysis and records sample size, estimate, uncertainty, and significance. These checks show which conclusions persisted and which depended on corpus construction or analytical settings.

Table 8. sensitivity and robustness checks

Sensitivity condition	Records or setting	Estimate	95% confidence interval	*p* value	Keyword map
Full corpus	44	IRR = 1.20	[1.08, 1.34]	.001	5 clusters
Scopus-indexed subset	34	IRR = 1.19	[1.05, 1.35]	.007	5 clusters
Web of Science-indexed subset	27	IRR = 1.17	[1.01, 1.37]	.041	4 clusters
Conference papers excluded	36	IRR = 1.18	[1.05, 1.33]	.006	4 clusters
Citation-age-adjusted comparison	44	H = 4.91; ε^2 = .08	—	.086	—
Keyword threshold raised to 3 occurrences	44; 16 terms retained	—	—	—	4 clusters

The annual publication trend remained statistically significant in the database-specific subsets and after conference papers were excluded. The Scopus-indexed subset produced an incidence rate ratio of 1.19, while the Web of Science subset produced 1.17; these subsets overlapped and were not independent samples. After citation-age adjustment, the document-type comparison was not statistically significant (p = .086), and its effect size was smaller. Raising the keyword threshold reduced the map from five clusters to four. These checks suggest that the example publication trend was stable, while citation and thematic results depended more on analytical choices.

Tables, Figures, and Results Summary

Table 9 consolidates the quantitative results reported across the corpus, productivity, citation, collaboration, and thematic analyses. It provides a compact cross-reference to the chapter's main numerical findings, including sample size, dominant categories, network structure, and robustness. Values remain examples and are not verified outputs from database screening or analysis.

Table 9. summary of principal quantitative findings

Results domain	finding	Numerical result
Final corpus	Publications included	44
Publication period	Years represented	2015–2025
Document composition	Journal articles	30 (68.2%)
Authorship	Multi-authored publications	36 (81.8%)
Scholarly influence	Total citations	612
Citation distribution	Median citations per publication	8 (IQR 4–20)
Publication trend	Annual incidence rate ratio	1.20 (95% CI [1.08, 1.34], p = .001)
Author collaboration	Nodes, links, clusters	38, 42, 5
Reference co-citation	Nodes, links, clusters	29, 74, 4
Keyword co-occurrence	Retained terms, links, clusters	24, 81, 5
Citation-age sensitivity	Document-type comparison	H = 4.91, ε^2 = .08, p = .086

The results summarized a 44-publication corpus covering 2015–2025. Journal articles represented 68.2% of records, and multi-authored papers represented 81.8%. The corpus accumulated 612 citations, with a median of eight citations per publication. The example trend model indicated increasing annual output. Network analyses showed five author-collaboration clusters and four reference co-citation clusters, while keyword mapping produced five clusters. The citation-age-adjusted comparison was not statistically significant. Together, these values provide a concise overview of productivity, influence, collaboration, and themes, while keeping descriptive indicators separate from direct measures of research quality or AI effectiveness.

Table 10 links each figure to the data series or network summary it represented. The values specify the number of years, records, nodes, links, or clusters displayed, supporting consistent captions and interpretation.

It summarizes the visuals accompanying the results and connects each figure with its corresponding quantitative analysis.

Table 10. figure content and quantitative inputs

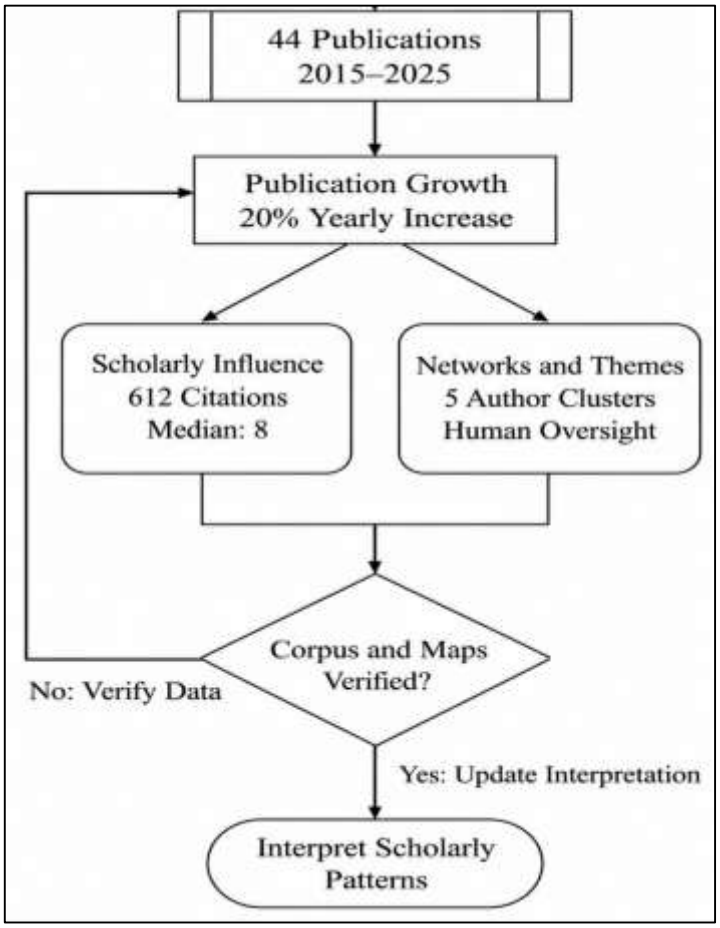
Figure	Visual content	values displayed
Figure 1	Annual publication trend	11 years; 44 publications; peak of 7 in 2023 and 2024
Figure 2	Leading authors, institutions, countries, and sources	4 contributor categories; top country count of 12; top source count of 5
Figure 3	Citation distribution	44 publications; 612 citations; median 8; maximum 66
Figure 4	Author co-authorship network	38 nodes; 42 links; 5 clusters
Figure 5	Citation-based intellectual structure	Co-citation: 29 nodes, 74 links, 4 clusters; coupling: 26 nodes, 48 links, 4 clusters
Figure 6	Keyword and thematic map	24 terms; 81 links; 5 clusters
Figure 7	Sensitivity comparison	3 keyword/network specifications; 4–5 clusters

The figure set represented publication output, contributor rankings, citation patterns, collaboration, intellectual structure, and themes. Figure 1 covered 11 publication years, while Figure 3 summarized citation counts for all 44 records. Figures 4–6 displayed network structures with node, link, and cluster counts reported in their captions. Figure 7 compared alternative mapping specifications, under which the number of clusters ranged from four to five. These numerical inputs made the figures traceable to the tables and analytical settings. In the completed chapter, each caption would state the database corpus, counting method, threshold, and retrieval or analysis date.

DISCUSSION

The analysis described a 44-publication corpus spanning 2015–2025. Annual publication output increased over the period, with the example model estimating a 20% yearly rise. Counts increased from one publication in 2015 to seven in 2023 and 2024, before declining to six in 2025. This pattern was broadly consistent with earlier bibliometric work describing the expansion of research on artificial intelligence, automation, and digital transformation in accounting and auditing (Danvila-del-Valle et al., 2019). However, the trend represented indexed publications retrieved under a specific search strategy; it did not measure real-world adoption, research quality, or practical impact. The smaller count in 2025 could also reflect indexing delays. Journal articles represented 30 of the 44 example records, and multi-authored publications accounted for 81.8%. The United States had the highest country count, followed by the United Kingdom, Malaysia, China, and Australia. These results suggested an internationally distributed but uneven publication profile (Aria et al., 2020). Earlier reviews of AI in accounting have also mapped publication activity across countries, institutions, and disciplines. Such comparisons help position reconciliation research within the wider field, but country rankings depend on database coverage, language limits, affiliation coding, and counting rules. Full counting allowed a multi-country or multi-author publication to contribute to several categories, so the percentages were not mutually exclusive. The example corpus accumulated 612 citations, with a median of eight per publication. Citation counts were concentrated: the most-cited records received substantially more attention than the typical publication (Li et al., 2018). The initial comparison suggested differences in citation distributions across journal articles, conference papers, and reviews. After adjusting for publication age, however, the comparison was no longer statistically significant. This change illustrated why citation totals should be interpreted alongside publication dates and document types. Earlier bibliometric studies use citations to identify influential works and intellectual foundations, but citation volume reflects scholarly attention rather than methodological quality or effectiveness. The example results therefore described productivity and influence without treating them as direct evidence of the value or accuracy of AI-assisted reconciliation.

Figure 11: Interpreting Bibliometric Research Findings



The network analysis identified five author co-authorship clusters, four institutional clusters, and four country clusters. Co-citation and bibliographic-coupling analyses also showed connected but distinct intellectual structures (Shen et al., 2019). These patterns were consistent with earlier science-mapping studies that used collaboration networks to describe research communities and citation networks to identify shared foundations. They did not establish the quality or impact of connected publications. Network results depended on inclusion thresholds, name normalization, and counting methods, while centrality and link strength described structural position rather than conventional statistical significance. The keyword analysis grouped publications around AI and accounting automation, reconciliation and financial close, audit analytics and internal control, human oversight and explainability, and transparency and reporting quality. Earlier bibliometric reviews of AI in accounting have commonly emphasized automation, machine learning, and audit analytics; related reporting studies have considered internal control and transparency. The example map brought these subjects into one framework, while showing that technical and human-centered themes may remain partly separated (Dai et al., 2020). Its structure changed when keyword thresholds or harmonization rules changed, illustrating that thematic maps depend on analytical choices and indexing language. A term's absence from a map would not necessarily mean that the subject was absent from the literature. Overall, the findings portrayed a focused research area within broader scholarship on accounting technology and financial reporting. Publication trends described output; citations and networks described scholarly attention and connections; keywords represented conceptual groupings. These indicators did not establish whether AI improved matching accuracy, reviewer decisions, internal controls, or reporting transparency. Those outcomes require examination of individual studies and their measures. The central interpretive point was that technical performance, human review, control effectiveness, and transparency are related but distinct. A high match rate alone does not show that exceptions were resolved appropriately or that supporting evidence became clearer (Khalaki et al., 2021). The discussion

remains provisional because the counts and estimates were examples rather than results from a completed search and analysis. It should be updated after the screened corpus, citation data, and network maps have been verified.

CONCLUSION

This study examined the scholarly landscape surrounding human-in-the-loop artificial intelligence for account reconciliation and financial-reporting transparency through a quantitative bibliometric and science-mapping framework. In the results, the corpus contained 44 publications from 2015 to 2025, with publication output increasing across most of the period. Journal articles and multi-authored publications formed the largest portions of the example dataset, while citation attention was concentrated among a smaller group of records. Collaboration and citation networks formed multiple clusters, and keyword mapping grouped the literature around accounting automation, reconciliation and financial close, audit analytics, human oversight, explainability, internal control, and reporting transparency. These patterns suggested that the topic connected several research communities rather than forming a single, unified field. The example analyses also showed why publication counts, citation indicators, and network positions should be interpreted as measures of scholarly activity and structure, not as direct evidence that AI improves accounting outcomes. Technical performance, human review, control effectiveness, and financial-reporting transparency remained distinct dimensions that require careful operational definitions and appropriately designed empirical evaluation. Sensitivity checks further illustrated that some conclusions could depend on database coverage, publication age, document type, and keyword thresholds. The example publication trend remained relatively stable across alternative specifications, whereas the citation comparison weakened after age adjustment and the thematic map changed under alternative thresholds. These findings are provisional because the numerical corpus and analytical outputs were rather than generated from a completed database search. Accordingly, this text serves as a conclusion model, not a report of verified bibliometric evidence. A final conclusion requires confirmed search records, documented screening decisions, validated citation data, and reproducible network analyses. Once verified, the mapping can provide a structured account of publication growth, influential research, collaboration patterns, and thematic relationships at the intersection of AI-supported reconciliation and transparent reporting. It can also clarify how the literature connects computational matching and exception detection with human judgment, internal control, and reporting outcomes. The study's central contribution is therefore its integrated quantitative framework for examining those connected dimensions while preserving the distinction between scholarly patterns and demonstrated operational effects.

RECOMMENDATION

Research on human-in-the-loop artificial intelligence for account reconciliation and financial-reporting transparency should use clearly defined search strategies, eligibility criteria, and database coverage so that bibliometric results can be reproduced and compared. The final study should document its search strings, search dates, screening decisions, duplicate-removal procedures, and metadata-cleaning rules. Scopus and Web of Science records should be checked for overlap, and citation counts should be reported with retrieval dates and appropriate consideration of publication age. Keyword harmonization rules and network thresholds should also be stated because they can change the number and structure of thematic clusters. Bibliometric indicators should be paired with structured content coding that records each study's AI task, data source, model family, validation design, unit of analysis, and outcome measures. Technical performance should be reported separately from human-review behavior, internal-control outcomes, and financial-reporting transparency. Reconciliation studies should specify whether they examine transaction matching, account-period certification, exception investigation, or the full close workflow, and should define denominators for accuracy, exception rates, review rates, and completion. Evaluations should include false matches and missed matches alongside overall accuracy, since these errors can have different accounting consequences. Human-in-the-loop research should report reviewer roles, evidence access, recommendation acceptance, overrides, escalations, review time, and decision authority. Where feasible, these behavioral measures should be linked to the correctness and documentation of the final accounting decision. Studies assessing transparency should distinguish source-to-report traceability, evidence completeness, explanation clarity, disclosure quality, and audit-trail retention rather than treating them as interchangeable

outcomes. Comparative evaluations should use appropriate manual or rule-based baselines, report uncertainty and effect sizes, and account for organization size, data quality, transaction complexity, staffing, and system integration. Sensitivity analyses should test alternative database sets, citation-age adjustments, document-type restrictions, and keyword thresholds. These recommendations are especially relevant because the numerical examples presented in this draft have not been verified against a completed corpus. The final recommendations should therefore be aligned with confirmed bibliometric patterns and the actual methods and limitations observed in the included studies.

LIMITATIONS

The interpretation of this study is subject to limitations associated with bibliometric research and the provisional nature of the numerical examples presented in the preceding sections. The corpus figures, citation counts, trend estimates, and network statistics were and were not generated from a completed database search; they therefore cannot be treated as empirical findings. A final account will depend on the verified search results, screening decisions, and reproducible analyses. Even with a completed search, restricting retrieval to selected databases may omit relevant publications indexed elsewhere, including regional journals, conference proceedings, books, and research in adjacent disciplines. A language restriction can underrepresent scholarship published in other languages, while the chosen publication period and document types can shape the apparent development of the field. Search terms also create a coverage limitation: studies may address reconciliation, financial close, accounting automation, auditability, or human–AI collaboration without using the exact terms selected for the query. Conversely, broad terms may retrieve records with only a limited connection to the study topic. Bibliometric indicators have further limitations. Citation counts vary with publication age, database coverage, document type, and disciplinary citation practices, and they measure attention rather than research quality or practical effectiveness. Author and institution name variations can affect productivity and collaboration rankings, while full-counting conventions can attribute the same publication to multiple contributors or countries. Keyword maps depend on author indexing, synonym harmonization, minimum occurrence thresholds, and network algorithms; different choices may produce different clusters.

REFERENCES

[1]. Abbasi Khalaki, M., Moameri, M., Asgari Lajayer, B., & Astatkie, T. (2021). Influence of nano-priming on seed germination and plant growth of forage and medicinal plants. *Plant growth regulation*, *93*(1), 13-28.

[2]. Adhikari, A., Bansal, M., & Kumar, A. (2021). IFRS convergence and accounting quality: India a case study. *Journal of International Accounting, Auditing and Taxation*, *45*, 100430.

[3]. Aditya, D., & Ahmed, H. C. H. R. (2022). Decision Support Systems for Last-Mile Logistics Optimization: A Data-Driven Framework for e-Commerce Fulfillment Networks. *International Journal of Scientific Interdisciplinary Research*, *12*(08), 01–49. https://doi.org/10.63125/h46vp562

[4]. Adnan, S. (2024). Human Trust in AI-Assisted Project Decision Systems: A Mixed-Methods Analysis of Explainability, Adoption, and Managerial Decision Quality. *Journal of Sustainable Development and Policy*, *3*(04), 287-332. https://doi.org/10.63125/9c139h44

[5]. Ahmad, N., Ullah, Z., Arshad, M. Z., waqas Kamran, H., Scholz, M., & Han, H. (2021). Relationship between corporate social responsibility at the micro-level and environmental performance: The mediating role of employee pro-environmental behavior and the moderating role of gender. *Sustainable production and consumption*, *27*, 1138-1148.

[6]. Aksoy, T., & Gurol, B. (2021). Artificial intelligence in computer-aided auditing techniques and technologies (CAATTs) and an application proposal for auditors. In *Auditing ecosystem and strategic accounting in the digital era: Global approaches and new opportunities* (pp. 361-384). Springer.

[7]. Akter, S., Michael, K., Uddin, M. R., McCarthy, G., & Rahman, M. (2022). Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*, *308*(1), 7-39.

[8]. Aminul Islam, M. (2024). Interpreting Cyber-Threat Indicators in Enterprise Memory Forensics: A Mixed-Methods Study of Machine Learning Performance and Explainability. *Review of Applied Science and Technology*, *14*(12), 01–13. https://doi.org/10.63125/2p8hte92

[9]. Aminul Islam, M., & Md, A. (2023). Adaptive Class-Imbalance Learning for High-Sensitivity Fraud Detection in Large-Scale Digital Transaction Networks. *International Journal of Business and Economics Insights*, *13*(09), 01-33. https://doi.org/10.63125/bmxwv741

[10]. Amiram, D., Bozanic, Z., Cox, J. D., Dupont, Q., Karpoff, J. M., & Sloan, R. (2018). Financial reporting fraud and other forms of misconduct: a multidisciplinary review of the literature. *Review of Accounting Studies*, *23*(2), 732-783.

[11]. Amiri, S., Jafari Jozani, M., & Modarres, R. (2018). Randomly selected order statistics in ranked set sampling: A less expensive comparable alternative to simple random sampling. *Environmental and Ecological Statistics*, *25*(2), 237-256.

[12]. Anyon, Y., Roscoe, J., Bender, K., Kennedy, H., Dechants, J., Begun, S., & Gallager, C. (2019). Reconciling adaptation and fidelity: implications for scaling up high quality youth programs. *The journal of primary prevention*, *40*(1), 35-49.

[13]. Aria, M., Misuraca, M., & Spano, M. (2020). Mapping the Evolution of Social Research and Data Science on 30 Years of Social Indicators Research: M. Aria et al. *Social Indicators Research*, *149*(3), 803-831.
[14]. Asheke, M. (2023). Machine Learning–Based Predictive Quality Modeling for Defect Detection and Process Optimization in Advanced Manufacturing Systems. *Review of Applied Science and Technology*, *13*(12), 01–49. https://doi.org/10.63125/thy73b16
[15]. Asheke, M., & Md Maksudul, H. (2022). AI-Driven Demand Forecasting for Sporting-Goods and Teamwear Supply Chains Under Seasonal and Tournament Demand. *International Journal of Business and Economics Insights*, *2*(1), 73–110. https://doi.org/10.63125/ed6by557
[16]. Asheke, M., & Mohammad Badrul, A. (2022). Development Of Functionalized Cellulose Nanofiber Adsorbents for Selective Critical-Metal Recovery from Industrial Wastewater. *American Journal of Scholarly Research and Innovation*, *1*(02), 270–320. https://doi.org/10.63125/v6fxzq09
[17]. Ashik, M. (2024). Meta-Analysis of Risk-Based Inspection Strategies for Improving Reliability and Safety Across Critical Energy Infrastructure Systems. *American Journal of Advanced Technology and Engineering Solutions*, *14*(09), 01-43. https://doi.org/10.63125/tgttkt09
[18]. Avvisati, F., Le Donne, N., & Paccagnella, M. (2019). A meeting report: cross-cultural comparability of questionnaire measures in large-scale international surveys. *Measurement instruments for the Social Sciences*, *1*(1), 8.
[19]. Baughman, A. W., Triantafylidis, L. K., Neil, N. O., Norstrom, J., Okpara, K., Ruopp, M. D., Linsky, A., Schnipper, J., Mixon, A. S., & Simon, S. R. (2021). Improving Medication Reconciliation with Comprehensive Evaluation at a Veterans Affairs Skilled-Nursing Facility. *The Joint Commission Journal on Quality and Patient Safety*.
[20]. Bennett, N. J., Kaplan-Hallam, M., Augustine, G., Ban, N., Belhabib, D., Brueckner-Irwin, I., Charles, A., Couture, J., Eger, S., & Fanning, L. (2018). Coastal and Indigenous community access to marine resources and the ocean: A policy imperative for Canada. *Marine Policy*, *87*, 186-193.
[21]. Bernstein, R. S., Bulger, M., Salipante, P., & Weisinger, J. Y. (2020). From Diversity to Inclusion to Equity: A Theory of Generative Interactions: RS Bernstein et al. *Journal of business ethics*, *167*(3), 395-410.
[22]. Brown, D. C., Lariscy, J. T., & Kalousová, L. (2019). Comparability of Mortality Estimates from Social Surveys and Vital Statistics Data in the United States: DC Brown et al. *Population research and policy review*, *38*(3), 371-401.
[23]. Buchholtz, G. (2019). Artificial intelligence and legal tech: Challenges to the rule of law. In *Regulating artificial intelligence* (pp. 175-198). Springer.
[24]. Bush, J. T. (2020). Win-Win-Lose? Sustainable HRM and the promotion of unsustainable employee outcomes. *Human Resource Management Review*, *30*(3), 100676.
[25]. Cao, Q.-T., Vuong, Q.-H., Pham, H.-H., Luong, D.-H., Ho, M.-T., Hoang, A.-D., & Do, M.-T. (2021). A bibliometric review of research on international students' mental health: Science mapping of the literature from 1957 to 2020. *European journal of investigation in health, psychology and education*, *11*(3), 781-794.
[26]. Cao, S., Powell, W., Foth, M., Natanelov, V., Miller, T., & Dulleck, U. (2021). Strengthening consumer trust in beef supply chain traceability with a blockchain-based human-machine reconcile mechanism. *Computers and electronics in agriculture*, *180*, 105886.
[27]. Chiewchantanakit, D., Meakchai, A., Pituchaturont, N., Dilokthornsakul, P., & Dhippayom, T. (2020). The effectiveness of medication reconciliation to prevent medication error: A systematic review and meta-analysis. *Research in Social and Administrative Pharmacy*, *16*(7), 886-894.
[28]. Dai, Y., Zheng, H., Jiang, Z., & Xing, B. (2020). Combined effects of biochar properties and soil conditions on plant growth: A meta-analysis. *Science of the total environment*, *713*, 136635.
[29]. Danvila-del-Valle, I., Estévez-Mendoza, C., & Lara, F. J. (2019). Human resources training: A bibliometric analysis. *Journal of Business Research*, *101*, 627-636.
[30]. de Fine Licht, K., & de Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision-making: Why explanations are key when trying to produce perceived legitimacy. *AI & society*, *35*(4), 917-926.
[31]. De Roeck, K., & Farooq, O. (2018). Corporate social responsibility and ethical leadership: Investigating their interactive effect on employees' socially responsible behaviors. *Journal of business ethics*, *151*(4), 923-939.
[32]. Doz, Y. (2020). Fostering strategic agility: How individual executives and human resource practices contribute. *Human Resource Management Review*, *30*(1), 100693.
[33]. Drydakis, N. (2022). Artificial Intelligence and reduced SMEs' business risks. A dynamic capabilities analysis during the COVID-19 pandemic. *Information Systems Frontiers*, *24*(4), 1223-1247.
[34]. Escarcha, J. F., Lassa, J. A., & Zander, K. K. (2018). Livestock under climate change: a systematic review of impacts and adaptation. *Climate*, *6*(3), 54.
[35]. Fernández-Macías, E., & Bisello, M. (2022). A comprehensive taxonomy of tasks for assessing the impact of new technologies on work. *Social Indicators Research*, *159*(2), 821-841.
[36]. Filgueiras, F. (2022). New Pythias of public administration: ambiguity and choice in AI systems as challenges for governance. *AI & society*, *37*(4), 1473-1486.
[37]. Flores-Romero, M. B., Perez-Romero, M. E., Alvarez-Garcia, J., & del Río-Rama, M. d. l. C. (2021). Bibliometric mapping of research on magic towns of Mexico. *Land*, *10*(8), 852.
[38]. Gbededo, M. A., Liyanage, K., & Garza-Reyes, J. A. (2018). Towards a Life Cycle Sustainability Analysis: A systematic review of approaches to sustainable manufacturing. *Journal of Cleaner Production*, *184*, 1002-1015.
[39]. Gil-Garcia, J. R., Guler, A., Pardo, T. A., & Burke, G. B. (2019). Characterizing the importance of clarity of roles and responsibilities in government inter-organizational collaboration and information sharing initiatives. *Government Information Quarterly*, *36*(4), 101393.

[40]. Gnewuch, U., Morana, S., Adam, M. T., & Maedche, A. (2022). Opposing Effects of Response Time in Human–Chatbot Interaction: U. Gnewuch et al. *Business & Information Systems Engineering*, *64*(6), 773-791.
[41]. Goode, J.-V., Owen, J., Page, A., & Gatewood, S. (2019). Community-based pharmacy practice innovation and the role of the community-based pharmacist practitioner in the United States. *Pharmacy*, *7*(3), 106.
[42]. Hermann, E. (2022). Leveraging artificial intelligence in marketing for social good—An ethical perspective. *Journal of business ethics*, *179*(1), 43-61.
[43]. Jagannathan, K., Arnott, J. C., Wyborn, C., Klenk, N., Mach, K. J., Moss, R. H., & Sjostrom, K. D. (2020). Great expectations? Reconciling the aspiration, outcome, and possibility of co-production. *Current Opinion in Environmental Sustainability*, *42*, 22-29.
[44]. Konigsburg, J. A. (2022). Modern warfare, spiritual health, and the role of artificial intelligence. *Religions*, *13*(4), 343.
[45]. Kovalerchuk, B., Andonie, R., Datia, N., Nazemi, K., & Banissi, E. (2022). Visual knowledge discovery with artificial intelligence: Challenges and future directions. In *Integrating artificial intelligence and visualization for visual knowledge discovery* (pp. 1-27). Springer.
[46]. Kynatun, M., & Saba, A. (2024). Machine Learning Enhanced Intrusion Detection and Anomaly Classification for MIS-Driven Enterprise Network Security. *International Journal of Scientific Interdisciplinary Research*, *14*(09), 01–41. https://doi.org/10.63125/xk6kpc72
[47]. Lefèvre, T., Colineaux, H., Morgand, C., Tournois, L., Delpierre, C., Delpierre, C., Lefèvre, T., Delpierre, C., Lefèvre, T., & Lefèvre, T. (2022). AI, Epidemiology and Public Health in the Covid Pandemic. In *Artificial Intelligence in Covid-19* (pp. 279-338). Springer.
[48]. Li, J. J., Bonn, M. A., & Ye, B. H. (2019). Hotel employee's artificial intelligence and robotics awareness and its impact on turnover intention: The moderating roles of perceived organizational support and competitive psychological climate. *Tourism management*, *73*, 172-181.
[49]. Li, K. X., Jin, M., & Shi, W. (2018). Tourism as an important impetus to promoting economic growth: A critical review. *Tourism management perspectives*, *26*, 135-142.
[50]. Li, S.-A., Jeffs, L., Barwick, M., & Stevens, B. (2018). Organizational contextual features that influence the implementation of evidence-based practices across healthcare settings: a systematic integrative review. *Systematic reviews*, *7*(1), 72.
[51]. López-Robles, J. R., Cobo, M. J., Gamboa-Rosales, N. K., & Herrera-Viedma, E. (2020). Mapping the intellectual structure of the international journal of computers communications and control: A content analysis from 2015 to 2019. International Conference on Computers Communications and Control,
[52]. López Belmonte, J., Segura-Robles, A., Moreno-Guerrero, A.-J., & Parra-González, M. E. (2020). Machine learning and big data in the impact literature. A bibliometric review with scientific mapping in Web of Science. *Symmetry*, *12*(4), 495.
[53]. Malenfant, S., Jaggi, P., Hayden, K. A., & Sinclair, S. (2022). Compassion in healthcare: an updated scoping review of the literature. *BMC palliative care*, *21*(1), 80.
[54]. Manning, L., Brewer, S., Craigon, P. J., Frey, J., Gutierrez, A., Jacobs, N., Kanza, S., Munday, S., Sacks, J., & Pearson, S. (2022). Artificial intelligence and ethics within the food sector: Developing a common language for technology adoption across the supply chain. *Trends in Food Science & Technology*, *125*, 33-42.
[55]. Mansura Akter, E., & Sumaiya, M. (2022). Antimicrobial Resistance Profiles of Acinetobacter Baumannii and Vibrio Cholerae in Hospital Wastewater and Adjacent Household Water Sources. *American Journal of Data Science and Analytics*, *3*(09), 01-55. https://doi.org/10.63125/gbervt14
[56]. Martens, W., Yapa, P. W., & Safari, M. (2020). The impact of financial statement comparability on earnings management: Evidence from frontier markets. *International Journal of Financial Studies*, *8*(4), 73.
[57]. Martinez, S., del Mar Delgado, M., Marin, R. M., & Alvarez, S. (2019). Science mapping on the Environmental Footprint: A scientometric analysis-based review. *Ecological Indicators*, *106*, 105543.
[58]. McCallig, J., Robb, A., & Rohde, F. (2019). Establishing the representational faithfulness of financial accounting information using multiparty security, network analysis and a blockchain. *International Journal of Accounting Information Systems*, *33*, 47-58.
[59]. Md Imran, H., & Jannatul, F. (2022). Molecular and Precision Rehabilitation Molecular Mechanisms Linking Mitochondrial Dysfunction, Inflammation, and Muscle Degeneration with Functional Decline and Rehabilitation Response in Neuromuscular Disease. *American Journal of Interdisciplinary Studies*, *12*(08), 01-44. https://doi.org/10.63125/0ym5bh79
[60]. Md Maksudul, H. (2024). Supplier-Risk Scoring and Procurement Resilience for Distributed Athletic-Apparel and Equipment Supply Networks. *American Journal of Advanced Technology and Engineering Solutions*, *14*(09), 44-90. https://doi.org/10.63125/a9fmgj50
[61]. Md Maksudul, H., & Md Tazul Islam, C. (2022). Machine Learning-Based Cash-Flow Forecasting for Improving Financial Planning in U.S. Small and Mid-Sized Businesses. *International Journal of Scientific Interdisciplinary Research*, *12*(12), 01–38. https://doi.org/10.63125/w3pgjx97
[62]. Md Shahid, N. (2023). Digital Twin-Based Network Reliability Modeling for LAN And WAN Infrastructure Resilience. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 577–629. https://doi.org/10.63125/ahjvxj14
[63]. Md Shahid, N. (2024). Predictive Analytics Models for Network Traffic Optimization in Distributed Enterprise Environments. *Review of Applied Science and Technology*, *3*(01), 367-419. https://doi.org/10.63125/qm6cyf84
[64]. Md Tazul Islam, C. (2024). Explainable AI For Financial and Operational Decision Support: Improving Transparency and Managerial Trust. *American Journal of Advanced Technology and Engineering Solutions*, *14*(12), 01-38. https://doi.org/10.63125/kwfvzw68

[65]. Md Tazul Islam, C., & Ahmed, H. C. H. R. (2023). Predictive Operations Analytics for Resource Allocation and Service-Capacity Planning in Data-Driven Organizations. *Review of Applied Science and Technology*, *13*(02), 01–41. https://doi.org/10.63125/z3sm3089

[66]. Md. Mainuddin, F. (2024). Quantitative Structural Retrofit Assessment Models for Strengthening Existing Steel Buildings Under Increased Load Demands. *Review of Applied Science and Technology*, *3*(04), 325–366. https://doi.org/10.63125/yyqnte84

[67]. Md. Morshedul, I., & Aminul Islam, M. (2022). A Quantitative Evaluation of Explainable Ensemble Learning for Probabilistic Risk Estimation and Decision Reliability in Enterprise Analytics. *American Journal of Data Science and Analytics*, *3*(02), 01-38. https://doi.org/10.63125/k22shk74

[68]. Melymuk, L., Nizzetto, P. B., Harner, T., White, K. B., Wang, X., Tominaga, M. Y., He, J., Li, J., Ma, J., & Ma, W.-L. (2021). Global intercomparison of polyurethane foam passive air samplers evaluating sources of variability in SVOC measurements. *Environmental science & policy*, *125*, 1-9.

[69]. Mengist, W., Soromessa, T., & Legese, G. (2020). Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX*, *7*, 100777.

[70]. Meshram, V. V., & Arora, J. (2021). Accounting constructs and economic consequences of IFRS adoption in India. *Journal of International Accounting, Auditing and Taxation*, *45*, 100427.

[71]. Mhlanga, D. (2022). Human-centered artificial intelligence: The superlative approach to achieve sustainable development goals in the fourth industrial revolution. *Sustainability*, *14*(13), 7804.

[72]. Nedd, R., Light, K., Owens, M., James, N., Johnson, E., & Anandhi, A. (2021). A synthesis of land use/land cover studies: Definitions, classification systems, meta-studies, challenges and knowledge gaps on a global landscape. *Land*, *10*(9), 994.

[73]. Nurunnabi, M., Donker, H., & Jermakowicz, E. K. (2022). The impact of mandatory adoption of IFRS in Saudi Arabia. *Journal of International Accounting, Auditing and Taxation*, *49*, 100509.

[74]. Nyanchoka, L., Tudur-Smith, C., Iversen, V., Tricco, A. C., & Porcher, R. (2019). A scoping review describes methods used to identify, prioritize and display gaps in health research. *Journal of Clinical Epidemiology*, *109*, 99-110.

[75]. Obwegeser, N., & Müller, S. D. (2018). Innovation and public procurement: Terminology, concepts, and applications. *Technovation*, *74*, 1-17.

[76]. Palaiogeorgou, P., Gizelis, C. A., Misargopoulos, A., Nikolopoulos-Gkamatsis, F., Kefalogiannis, M., & Christonasis, A. M. (2021). AI: Opportunities and challenges-The optimal exploitation of (telecom) corporate data. Conference on e-Business, e-Services and e-Society,

[77]. Paul, J., & Criado, A. R. (2020). The art of writing literature review: what do we know and what do we need to know? *International business review*, *29*(4), 101717.

[78]. Polak, P., Nelischer, C., Guo, H., & Robertson, D. C. (2020). "Intelligent" finance and treasury management: what we can expect. *AI & society*, *35*(3), 715-726.

[79]. Rashid, M. M., Askari, M. R., Chen, C., Liang, Y., Shu, K., & Cinar, A. (2022). Artificial intelligence algorithms for treatment of diabetes. *Algorithms*, *15*(9), 299.

[80]. Robotto, A., Quaglino, P., Lembo, D., Morello, M., Brizio, E., Bardi, L., & Civra, A. (2021). SARS-CoV-2 and indoor/outdoor air samples: a methodological approach to have consistent and comparable results. *Environmental Research*, *195*, 110847.

[81]. Rungvivatjarus, T., Kuelbs, C. L., Miller, L., Perham, J., Sanderson, K., Billman, G., Rhee, K. E., & Fisher, E. S. (2020). Medication reconciliation improvement utilizing process redesign and clinical decision support. *The Joint Commission Journal on Quality and Patient Safety*, *46*(1), 27-36.

[82]. Sadia, A., & Md. Mosheur, R. (2023). Consumer Trust, Perceived Authenticity, and Personalized Digital Marketing: A Quantitative Decision-Science Study. *International Journal of Scientific Interdisciplinary Research*, *4*(2), 252–295. https://doi.org/10.63125/cajkqt60

[83]. Shen, N., Liao, H., Deng, R., & Wang, Q. (2019). Different types of environmental regulations and the heterogeneous influence on the environmental total factor productivity: empirical analysis of China's industry. *Journal of Cleaner Production*, *211*, 171-184.

[84]. Silver, S. D. (2021). Dynamics of Negative evaluations in the information exchange of interactive decision-making teams: Advancing the design of technology-augmented GDSS. *Information Systems Frontiers*, *23*(6), 1621-1642.

[85]. Stabak, D., & Md Shahid, N. (2022). Empirical Assessment of Uncertainty-Aware Real-Time Decision-Making for Reliable and Resilient Cyber-Physical Systems. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 432–477. https://doi.org/10.63125/q3r4kk16

[86]. Sumaiya, M. (2024). Hospital Wastewater as A Reservoir of Antimicrobial Resistance Genes: Implications for Public Health and Genomic Surveillance in The United States. *Review of Applied Science and Technology*, *3*(04), 448–508. https://doi.org/10.63125/ahvas976

[87]. Syed Feroz, A., & Md Mahamud Pervaz, P. (2022). Root-Cause Analysis and FMEA-Driven Reliability Improvement in Gas Production Facilities: A Multi-Year Empirical Study. *International Journal of Scientific Interdisciplinary Research*, *12*(10), 01–53. https://doi.org/10.63125/70msp878

[88]. Syed Feroz, A., & Shabbir, A. (2022). Finite Element-Based Thermomechanical Optimization of Pressure Components for Extending Industrial Boiler Service Life. *Journal of Sustainable Development and Policy*, *12*(12), 01-51. https://doi.org/10.63125/x24pht28

[89]. Tabrez, A., Luebbers, M. B., & Hayes, B. (2020). A survey of mental modeling techniques in human–robot teaming. *Current Robotics Reports*, *1*(4), 259-267.

[90]. Taher, M. A., Sikder, M. A., Ahmed, H., & Bashar, M. (2024). Integrating IOT and Big Data Analytics for Enhanced Supply Chain Performance in Industrial Engineering Sectors: A Cross-Market Study. *International Journal of Science and Engineering*, *Volume 1*, 1-14.

[91]. Tamburino, L., & Bravo, G. (2021). Reconciling a positive ecological balance with human development: A quantitative assessment. *Ecological Indicators*, *129*, 107973.

[92]. Tams, S., Legoux, R., & Léger, P.-M. (2018). Smartphone withdrawal creates stress: A moderated mediation model of nomophobia, social threat, and phone withdrawal context. *Computers in Human Behavior*, *81*, 1-9.

[93]. Tosun, J., Galanti, M. T., & Howlett, M. (2022). The significance of leadership in the evolution of policy styles: reconciling policy-making in the short and long term. *Politische Vierteljahresschrift*, *63*(2), 337-358.

[94]. Tsao, M. S., Kerr, K. M., Kockx, M., Beasley, M.-B., Borczuk, A. C., Botling, J., Bubendorf, L., Chirieac, L., Chen, G., & Chou, T.-Y. (2018). PD-L1 immunohistochemistry comparability study in real-life clinical samples: results of blueprint phase 2 project. *Journal of Thoracic Oncology*, *13*(9), 1302-1311.

[95]. Umapathy, A., Chamley, L. W., & James, J. L. (2020). Reconciling the distinct roles of angiogenic/anti-angiogenic factors in the placenta and maternal circulation of normal and pathological pregnancies. *Angiogenesis*, *23*(2), 105-117.

[96]. van der Worp, H., Loohuis, A. M., Flohil, I. L., Kollen, B. J., Wessels, N. J., & Blanker, M. H. (2020). Recruitment through media and general practitioners resulted in comparable samples in an RCT on incontinence. *Journal of Clinical Epidemiology*, *119*, 85-91.

[97]. Velez-Estevez, A., García-Sánchez, P., Moral-Muñoz, J. A., & Cobo, M. J. (2022). Why do papers from international collaborations get more citations? A bibliometric analysis of Library and Information Science papers. *Scientometrics*, *127*(12), 7517-7555.

[98]. Vinodkumar, M., & Anoop, A. (2020). Review on comparability of 'classical'and 'contemporary'research methods in the context of Ayurveda. *Journal of Ayurveda and integrative medicine*, *11*(4), 539-546.

[99]. Vo, T. T. D., Tuliao, K. V., & Chen, C.-W. (2022). Work motivation: The roles of individual needs and social conditions. *Behavioral Sciences*, *12*(2), 49.

[100]. Werner, M., Wiese, M., & Maas, A. (2021). Embedding process mining into financial statement audits. *International Journal of Accounting Information Systems*, *41*, 100514.

[101]. Zhang, Y., & Xin, J. (2019). Survivable deployments of optical sensor networks against multiple failures and disasters: A survey. *Sensors*, *19*(21), 4790.

[102]. Zhang, Y., Zheng, Y., Zhang, L., Xu, S., Liu, X., & Chen, W. (2021). A meta-analytic review of the consequences of servant leadership: The moderating roles of cultural factors. *Asia Pacific Journal of Management*, *38*(1), 371-400.

[103]. Zhang, Z., Genc, Y., Wang, D., Ahsen, M. E., & Fan, X. (2021). Effect of AI explanations on human perceptions of patient-facing AI-powered healthcare systems. *Journal of Medical Systems*, *45*(6), 64.

[104]. Zhu, P., Zhao, Y., Agarwal, S., Henry, J., & Zinkle, S. J. (2022). Toward accurate evaluation of bulk hardness from nanoindentation testing at low indent depths. *Materials & Design*, *213*, 110317.