

Explainable Artificial Intelligence in Financial Fraud Detection: A Systematic Review and FinTech-Oriented ADO–TCCM Meta-Framework for Trust, Governance, and Transparency

Devansh Gupta ¹ , Priyanka Chugh ¹ , Kiran Sood ^{2,3}  and Simon Grima ^{4,5,*} 

¹ University School of Business, Chandigarh University, Mohali 140413, India; 23bcm70319@cuchd.in (D.G.); priyanka.e14624@cumail.in (P.C.)

² Chitkara Business School, Chitkara University, Chandigarh 140401, India; kiran.sood@chitkara.edu.in

³ Women Researchers Council (WRC), Azerbaijan State University of Economics (UNEC), Baku AZ1001, Azerbaijan

⁴ Department of Insurance and Risk Management, Faculty of Economics, Management and Accountancy, University of Malta, MSD2080 Msida, Malta

⁵ Faculty of Economics and Social Sciences, University of Latvia, LV-1586 Riga, Latvia

* Correspondence: simon.grima@um.edu.mt

Abstract

Artificial intelligence (AI)-driven fraud detection systems in FinTech ecosystems increasingly face a governance tension between high predictive accuracy and limited regulatory transparency, a gap that existing reviews have not addressed through an integrated behavioural, technical, and institutional lens. This study synthesises 99 Scopus-indexed, ABDC-ranked journal articles (2015–2026) using PRISMA 2020 and the SPAR-4-SLR protocol, integrating the Theory of Planned Behaviour (TPB) within an Antecedents–Decisions–Outcomes (ADO) framework to examine organisational adoption of explainable AI (XAI) in financial fraud detection. Three antecedent clusters are identified: attitudinal (algorithmic complexity, model opacity, data imbalance), normative (regulatory compliance, ethical expectations), and control-based (technical self-efficacy, organisational readiness)—which drive decision mechanisms including post hoc interpretability tools (SHapley Additive exPlanations [SHAP], Local Interpretable Model-Agnostic Explanations [LIME]), ethical governance protocols, and human-in-the-loop oversight. These produce outcomes across precision (reduced false positives, improved decision accuracy), compliance (audit transparency, institutional legitimacy), and cognitive (user acceptance, procedural justice) dimensions. The study introduces the Stability–Transparency–Reliability (STR) model, which advances TPB, Socio-Technical Systems Theory, and the Dynamic Capabilities View by reframing XAI from a static interpretability output into a recursive governance capability, formalised through the concept of Interpretative Agility, with direct implications for financial institutions operating under the EU AI Act.

Keywords: explainable artificial intelligence; financial fraud detection; theory of planned behaviour; ADO meta-framework; algorithmic governance; stability–transparency–reliability (STR) model; human-in-the-loop; socio-technical systems; interpretative agency; systematic literature review

JEL Classification: G28; G29; M15



Academic Editor: Shyan-Ming Yuan

Received: 18 April 2026

Revised: 18 June 2026

Accepted: 26 June 2026

Published: 8 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The increasing integration of Artificial Intelligence (AI) into financial systems is transforming how fraud detection, risk assessment, and governance processes are conducted. Traditional rule-based systems are being replaced by data-driven models that can identify complex patterns in large-scale financial data [1–3]. In the past, financial governance and fraud detection processes were based on deterministic, rule-driven systems that emphasised linear transparency rather than predictive depth [4–6]. However, the rapid acceleration of Artificial Intelligence (AI) has fundamentally transformed the architecture of contemporary financial intelligence, potentially shifting the governance paradigm from a human-centric oversight model toward a complex socio-technical partnership. In this environment, AI agents increasingly serve as active entities in deliberations over creditworthiness and fraudulent intent [7–12]. In this study, FinTech is used as a foundational construct, an abbreviation of Financial Technology, and warrants further definition based on the plurality of meanings advanced in the contemporary literature [13], which conceptualises FinTech as the application of technology to traditional financial services, emphasising process-level transformation within established institutional settings. The United Nations Department of Economic and Social Affairs (UN DESA, 2023) [14] foregrounds the financial inclusion dimension, defining FinTech as a vehicle for extending digital financial services to underserved populations. IBM (2024) [15] adopts a technical-infrastructure perspective, framing FinTech as the deployment of software and data architectures that automate and optimise financial operations. A further stream of industry-oriented definitions characterises FinTech as the fusion of finance and technology, with particular emphasis on transforming the design, delivery, and user experience of financial products. For the purposes of this study, FinTech is understood as the application of emerging digital technologies—including artificial intelligence, machine learning, and data analytics—to transform the design, delivery, governance, and user experience of financial services, with particular emphasis on scalability, regulatory compliance, and organisational transparency [13,16]. This operationalisation is deliberately governance-oriented: it positions FinTech not merely as a technological infrastructure but as an institutionally embedded ecosystem in which human adoption behaviour, regulatory compliance intent, and organisational trust are as consequential as technical performance—which is precisely why the Theory of Planned Behaviour (TPB) is the primary theoretical lens of this review.

Within this digital FinTech ecosystem, integrating AI into fraud detection systems represents a critical shift toward automated, data-driven financial services. As digital banking, payment platforms, and RegTech solutions expand, the demand for transparent and auditable AI systems becomes increasingly important. Within these ecosystems, explainability has emerged as a foundational governance requirement, as developed in full in Section 4.

As AI continues to develop, it has transitioned from a computational tool to increasingly autonomous decision-support systems, which could significantly affect how financial managers and auditors make decisions about very risky investments [17–19]. As a result, companies are increasingly being urged to undergo an ontological transformation of corporate governance, shifting from a traditional model in which AI is treated as an isolated technology to one in which AI is used as a substantive component of the corporate governance process [4,9,20,21]. While there is considerable potential for AI to deliver substantial economic benefits by identifying illicit trade and improving risk assessment [1,3], an emerging tension exists between its predictive capabilities and the need for regulatory transparency—a tension this study specifically addresses.

Although state-of-the-art models such as Graph Neural Networks (GNNs) and Deep Neural Networks (DNNs) can achieve very high accuracy (over 90% in some cases) in

identifying interrelated patterns of fraudulent behaviour, they do so with little transparency [11,22–24]. This has led to what is referred to as a ‘black box’ issue, in which an epistemological gap exists between how a decision is made and how it can be understood by those involved [19,25]. When accuracy and accountability become disconnected in high-stakes financial environments, trust among stakeholders may decline, information asymmetries may become more prevalent, and systemic vulnerabilities may emerge [2,26,27]. From a technical standpoint, explainable artificial intelligence (XAI) methods such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) provide post hoc interpretability by attributing model outputs to input features. While SHAP leverages cooperative game theory to ensure consistency and global interpretability, LIME approximates local decision boundaries using surrogate models. These techniques are increasingly embedded within FinTech fraud detection systems to reconcile predictive accuracy with interpretability requirements, particularly in high-frequency transaction environments.

If financial institutions do not adopt “Interpretative Agility,” they risk encountering the “dark side” of digital transformation: declining investigative efficiency and an inability to meet the “Strict Necessity Test” imposed by emerging global regulatory bodies [28–31]. Moreover, the existing literature on XAI as it relates to financial fraud detection is fragmented and organised into three competing perspectives [32,33]. First, the technical perspective is centred on optimising the detection rate through advanced hybrid architectures [34–36]. Second, the behavioural perspective investigates cognitive dimensions of AI adoption, focusing on how perceived risk and “Innovation Resistance” prevent an investigator from acting on an AI-generated fraud flag [37–40]. Thirdly, the regulatory perspective emphasises the need for legal mandates on transparency, such as those outlined in the EU AI Act, but fails to provide the technical detail needed to operationalise them [7,24,41]. These three streams have developed largely in parallel rather than in dialogue with one another: technology-oriented research, typified by computer science and engineering venues, evaluates explainability primarily through computational benchmarks such as fidelity and simulatability, whereas governance-oriented research (the epistemological domain of the present review) treats interpretability as an institutional capability assessed against regulatory accountability, organisational trust, and audit readiness [25,28]. Critically, no consensus exists within the literature on what ‘interpretability’ means in practice. Certain studies distinguish simulatability, decomposability, and algorithmic transparency as distinct constructs, while certain studies frame interpretability as a spectrum from local post hoc explanations to global model transparency—without these perspectives converging on a common evaluative standard. This definitional plurality partly explains why the three research streams have produced incommensurable findings and why an integrative, behaviourally grounded framework is required.

The absence of an integrated framework linking technical, behavioural, and institutional silos is a significant gap in the existing foundational literature. Presently, only a limited number of studies demonstrate how technological Antecedents (e.g., stable feature engineering) can affect Organisational Outcomes, such as trust or compliance [19,42–44]. Additionally, there is a lack of studies exploring the “Reject Inference” issue in fraud modelling; as a result, the behaviour of excluded samples is often ignored, potentially leading to biased decisions [45]. Therefore, a multi-level model is sought to investigate the potential for XAI as a strategic, cognitive, and governance enabler within organisational decision-making systems.

Despite the growing body of systematic literature reviews on explainable artificial intelligence in finance, existing studies predominantly focus on either algorithmic performance or isolated governance dimensions, with limited integration across behavioural,

technical, and institutional perspectives. In particular, prior reviews have not explicitly incorporated behavioural theory to explain organisational adoption of explainable AI in high-stakes financial decision-making contexts. Furthermore, the methodological architecture of existing XAI studies is dominated by quantitative algorithmic experimentation ($\approx 60\%$ of the reviewed corpus; see Section 4.5), with longitudinal governance studies, qualitative institutional analyses, and cross-contextual replication studies remaining substantially underrepresented, a methodological imbalance that limits the generalisability of current findings beyond controlled benchmark datasets.

Based on the above discussion, three key research gaps emerge:

- (i) The lack of an integrated framework linking behavioural, technical, and institutional drivers of explainable AI adoption;
- (ii) Limited exploration of explainability as a governance capability within financial systems;
- (iii) Insufficient attention to the governance and capability-development dimensions of explainability within FinTech environments.

Accordingly, this study is guided by the following research questions:

RQ1: What are the antecedents, decision processes, and outcomes influencing explainable AI-driven financial fraud detection?

RQ2: What are the major geographical regions, journals of publication, and types of research articles in the explainable AI–financial fraud detection literature?

RQ3: What are the significant research gaps in the explainable AI–financial fraud detection domain?

RQ4: What are the theoretical and managerial implications, and what future research directions emerge from the explainable AI–financial fraud detection literature?

A hierarchical theoretical approach is adopted to address these questions, as detailed in the Section titled Theoretical Architecture: Hierarchy and Functional Role.

Theoretical Architecture: Hierarchy and Functional Role

The eight theories and two analytical protocols employed in this study are not co-equal constructs deployed in parallel. They operate in a deliberate three-tier hierarchy, as illustrated in Figure 1. At the apex, the Theory of Planned Behaviour (TPB) serves as the primary behavioural lens governing all three ADO layers—it is the organising logic of the entire framework. At the second tier, three theories provide layer-specific enrichment: the Resource-Based View (RBV) operationalises the Antecedents layer by conceptualising XAI as an evolving organisational capability; Decision Theory operationalises the Decisions layer by examining how interpretability tools reduce cognitive load and automation bias; and Institutional Theory operationalises the Outcomes layer by explaining how regulatory and social pressures drive legitimacy-seeking behaviour. The Dynamic Capabilities View (DCV) and Socio-Technical Systems (STS) Theory occupy a cross-cutting role, spanning all three layers: DCV provides the temporal dimension through a recursive learning loop, while STS Theory frames the co-evolution of human and technical subsystems within governed financial environments. At the base, ADO and TCCM are not substantive theories but analytical protocols—ADO structures the causal flow of the synthesis, and TCCM ensures its methodological rigour. For clarity, two of the analytical protocols depicted in this hierarchy warrant a brief definitional grounding prior to Figure 1. The ADO (Antecedents–Decisions–Outcomes) framework is a systematic analytical structure that traces the causal flow of a phenomenon (in this case, XAI adoption) from its behavioural and institutional drivers (Antecedents), through the mechanisms by which those drivers are operationalised (Decisions), to the organisational and societal impacts that result (Outcomes). The TCCM (Theory, Context, Characteristics, Methodology) protocol is a systematic overlay that ensures the synthesis is evaluated across four methodological dimensions: the theoretical

underpinnings of included studies, the contexts in which the findings were produced, the characteristics of the research designs, and the methodological approaches employed—together ensuring that the ADO synthesis is not merely descriptive but epistemologically structured. This architecture is depicted in Figure 1 below.

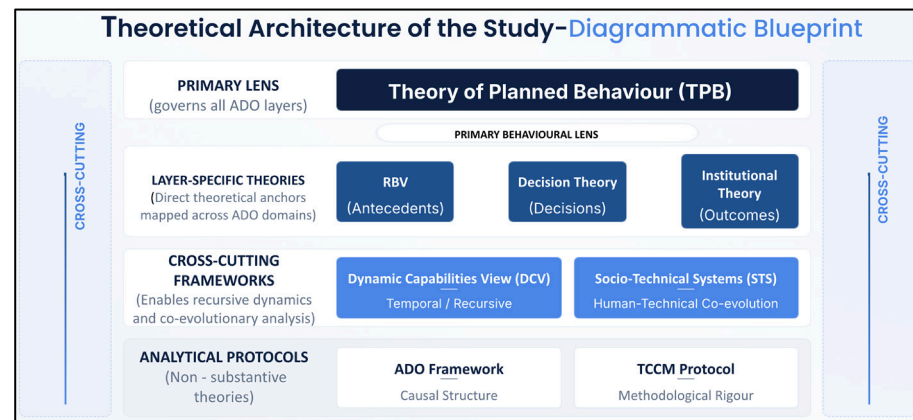


Figure 1. Theoretical Architecture of the Study: Hierarchy and Functional Role of Frameworks (Authors’ Compilation).

The primary theoretical lens is the Theory of Planned Behaviour (TPB), which lays the behavioural foundation for understanding the interplay of cognitive, normative and control factors in governing XAI adoption in fraud detection contexts. Supporting theories serve as layer-specific extensions rather than parallel frameworks. RBV conceptualises XAI as a continually evolving organisational capability within the “Antecedents” layer [46–48]. Institutional Theory explores how external pressures for transparency require organisations to recalibrate their “Outcomes” [49–51]. Decision Theory examines how specific XAI “Decisions” may mitigate cognitive load and “automation bias” [52–54]. The Dynamic Capabilities View (DCV) is employed to explore how XAI serves as a “Learning Catalyst” for organisations, thereby reconfiguring a firm’s future organisational readiness through a recursive feedback loop [31,55–57]. The ADO (Antecedents–Decisions–Outcomes) framework operationalises this hierarchy as the primary analytical structure, tracing the causal flow of explainability from behavioural intent to institutional outcomes. At the same time, the TCCM protocol (Theory, Context, Characteristics, Methodology) ensures methodological rigour in synthesising the 99-article corpus.

Finally, this study proposes the STR (Stability–Transparency–Reliability) model as an integrative framework that synthesises insights from the reviewed literature and extends existing theoretical perspectives on AI governance. The study proposes the concept of “Interpretative Agility”—an institutional capability that enables firms to maintain stable, causal justifications for fraud flags even during periods of market volatility or concept drift [28,58,59]. For conceptual clarity this study differentiates between the Interpretative Agency and the Interpretative Agility. Interpretative Agency is the cognitive ability of the individual or organisational actor to forensically validate and justify AI-generated outputs, an extension of TPB’s Perceived Behavioural Control construct. Interpretative Agility, on the other hand, is the institution’s dynamic capacity to continuously maintain, adapt and reproduce stable explanatory logic across changing operational environments, extending the Dynamic Capabilities View through the Recursive Learning Loop. By potentially facilitating a higher level of trust calibration, Interpretative Agility is designed to foster an effective human–AI dyad that is audit-ready and aligned with the “Right to Explanation” established by global privacy mandates [39,60,61].

The remainder of this paper is organised as follows: Section 2 details the research methodology, comprising the review design, pilot search strategy, operationalisation and inclusion–exclusion framework, and the use of PRISMA 2020 guidelines and SPAR-4-SLR Protocol. Section 3 presents the descriptive analysis, followed by the Antecedent–Decision–Outcome (ADO) Framework illustrated in Section 4. Further, Section 5 discusses the findings, followed by Implications (Theoretical and Managerial) and Future Research Directions in Sections 6 and 7, respectively.

Four primary contributions emerge from this study:

1. A FinTech-oriented framework is developed, integrating explainable AI into fraud detection and governance systems;
2. The ADO framework is extended by incorporating behavioural theory (TPB) within AI-driven financial decision-making;
3. The Stability–Transparency–Reliability (STR) model is introduced as a governance mechanism for explainability, operationalised through three named theoretical extensions: (a) Interpretative Agency, an extension of TPB’s Perceived Behavioural Control from resource availability into cognitive–forensic validation capacity; (b) Triadic STS Alignment, an extension of Socio-Technical Systems Theory incorporating a regulatory–ethical subsystem as a non-negotiable governance stabiliser; and (c) the Recursive Learning Loop, an extension of the Dynamic Capabilities View that formalises XAI as an active, self-reinforcing governance catalyst rather than a static compliance instrument, operationalised through the concept of Interpretative Agility. These extensions are developed in full in Section 6.1.
4. A systematic synthesis of XAI applications in financial fraud detection is provided, with implications for regulatory governance, practitioner decision-making, and future research directions in FinTech.

2. Research Methods

2.1. Review Design

A systematic literature review (SLR) methodology is adopted to assess the intellectual framework of Explainable Artificial Intelligence (XAI) for detecting financial fraud. The method uses the principles of evidence-based review design as laid down by Tranfield et al. (2003 [62]; Snyder, 2019 [63] and Webster & Watson, 2002 [64] to develop replication, transparency, and conceptual integrity within this area.

The authors developed a structured workflow for conducting the review, as illustrated in Figure 2, which included four stages: a pilot exploration stage; retrieving data from the Scopus database; screening for eligibility; and synthesising using PRISMA 2020 [65] and the SPAR-4-SLR methodology [66].

The combined use of PRISMA 2020 and the SPAR-4-SLR protocol enhances both methodological transparency and analytical depth. While PRISMA ensures systematic identification and screening of relevant studies, SPAR-4-SLR enables structured conceptual synthesis, making this approach particularly suitable for interdisciplinary domains such as FinTech, where technical, behavioural, and regulatory perspectives intersect. Also, a formal review protocol was not prepared prior to conducting this review; this systematic review was not prospectively registered in PROSPERO or any other review registry.

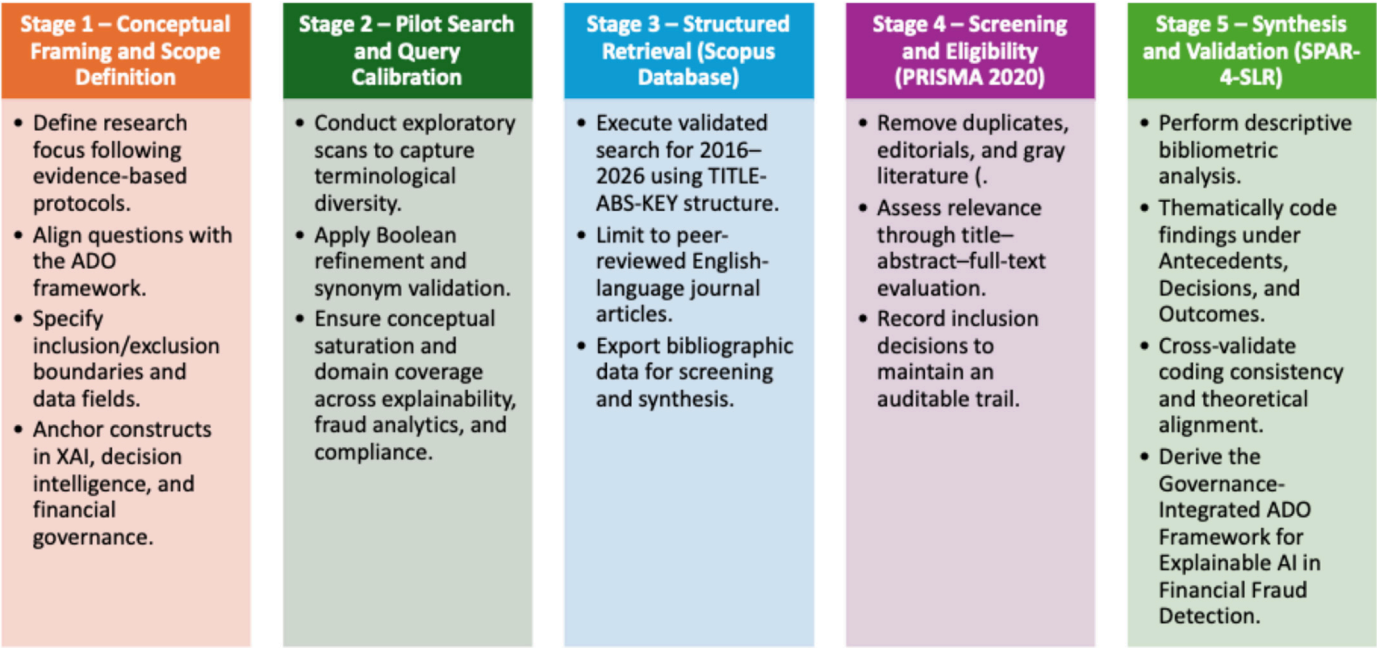


Figure 2. Systematic Review Framework.

2.2. Pilot Search and Query Calibration

An initial investigation into the terminology and boundaries of the field was conducted using the approach described by Okoli and Schabram 2010 [67] and Kitchenham and Charters, 2007 [68]. Keyword clustering with Boolean refinement was used to assess the completeness of conceptual constructs related to the Interpretability of AI, Responsible AI, and Financial Fraud Detection.

Iterative testing continued until both conceptual saturation and balance across disciplines were achieved. This ensured that the retrieved results from Scopus would represent the intersection of AI explainability, decision science, and governance, providing a linguistic foundation that is validated before conducting an integrated search. To enhance reproducibility, the full search strategy, inclusion criteria, and screening process are documented to allow future researchers to replicate or extend the review.

2.3. Operational Search Strategy and Eligibility Criteria

To retrieve data from Scopus database, a structured Boolean query that integrated three conceptual clusters of academic relevance, including AI-based computational systems, financial fraud and economic crime applications, and governance–transparency constructs, was devised and executed in Scopus Database on 20 February 2026. As detailed in Table 1, inclusion parameters restricted results to English-language, open-access, final-stage journal articles indexed under Business, Management and Accounting and Economics, Econometrics and Finance. Post-retrieval from the Scopus database, only articles published in journals listed in the Australian Business Deans Council (ABDC) Journal Quality were considered for analysis (as shown in Table 1).

Table 1. Inclusion and Exclusion Criteria and Operationalisation of the Scopus Search Strategy.

	Dimension	Inclusion Criteria	Exclusion Criteria	Operationalisation in Scopus Query
KEYWORD SEARCH STRING	Article Fields Searched	Studies where relevant concepts appear in the title, abstract, or author keywords to ensure thematic centrality	Studies where search terms appear only in full text, references, or Supplementary Material	TITLE-ABS-KEY
	Core Technological Focus: AI-Based Systems	Research explicitly applying artificial intelligence, machine learning, AI-driven systems, algorithmic or automated decision-making, predictive analytics, intelligent systems, or data-driven models.	Studies addressing fraud or risk without AI-based computational or algorithmic approaches	TITLE-ABS-KEY ("machine learning" OR "artificial intelligence" OR "AI-driven" OR "algorithmic decision*" OR "automated decision*" OR "predictive analytics" OR "intelligent system*" OR "data-driven model*")
	Application Domain: Financial Fraud and Economic Crime	Studies examining financial fraud detection/prevention, financial crime, economic crime, money laundering, AML, suspicious transactions, transaction monitoring, or financial risk assessment	AI studies unrelated to fraud, economic crime, or financial compliance contexts	AND (fraud* OR "financial fraud" OR "fraud detection" OR "fraud prevention" OR "financial crime" OR "economic crime" OR "money laundering" OR "AML" OR "suspicious transaction*" OR "risk assessment" OR "transaction monitoring")
	Governance, Transparency and Regulatory Orientation	Research incorporating explainability, interpretability, transparency, accountability, governance, ethical AI, responsible AI, compliance, regulatory oversight, trust, fairness, risk management, or organisational performance.	Purely technical fraud-detection studies lack governance, ethical, compliance, and transparency considerations.	AND (explainab* OR interpret* OR transparency OR accountability OR governance OR trust OR "risk management" OR compliance OR "regulatory compliance" OR "regulatory oversight" OR "ethical AI" OR "responsible AI" OR "model transparency" OR "decision transparency" OR "algorithmic fairness" OR "organisational performance")
ELIGIBILITY FILTERS	Publication Period	Articles published between 2015 and 2026 (inclusive), reflecting the modern evolution of AI-driven fraud governance	Articles published before 2015 or after 2026	AND PUBYEAR > 2014 AND PUBYEAR < 2027
	Subject Area Scope	Studies indexed under Business, Management and Accounting and Economics, Econometrics and Finance.	Studies classified exclusively under engineering, computer science, medicine, or unrelated disciplines.	AND (LIMIT-TO (SUBJAREA, "BUSI") OR LIMIT-TO (SUBJAREA, "ECON"))
	Document Type	Peer-reviewed journal research articles	Conference papers, book chapters, editorials, reviews, notes, letters, or errata	AND (LIMIT-TO (DOCTYPE, "ar"))

Table 1. Cont.

		Dimension	Inclusion Criteria	Exclusion Criteria	Operationalisation in Scopus Query
ELIGIBILITY FILTERS		Source Type	Journal publications to ensure scholarly rigour and standardised indexing	Non-journal sources, such as conference proceedings or trade publications	AND (LIMIT-TO (SRCTYPE, "j"))
		Access Type	Open Access articles to ensure transparency, reproducibility, and accessibility	Closed-access or subscription-restricted publications	AND (LIMIT-TO (OA, "all"))
		Publication Stage	Final, fully published articles	Articles in press, early access, or pre-publication stages	AND (LIMIT-TO (PUBSTAGE, "final"))
		Language	Publications written in English to maintain analytical consistency	Publications in languages other than English	AND (LIMIT-TO (LANGUAGE, "English"))
		Journal Quality Filter	Articles published in journals recognised in the Australian Business Deans Council (ABDC) Journal Quality List to ensure established scholarly credibility and ranking standards	Articles published in journals recognised in the Australian Business Deans Council (ABDC) Journal Quality List to ensure established scholarly credibility and ranking standards	Articles published in journals recognised in the Australian Business Deans Council (ABDC) Journal Quality List to ensure established scholarly credibility and ranking standards

Three methodological decisions warrant explicit justification prior to documenting the study's limitations. Scopus was selected as the primary database because it provides the most comprehensive and consistently structured indexing of the business, management, economics, and finance literature among multi-disciplinary databases, and it indexes all journals listed in the ABDC Journal Quality List—the secondary quality filter applied in this review [62,69]. Unlike Web of Science, which carries a natural science and citation-impact weighting, or IEEE Xplore, which prioritises engineering and computer science outputs, Scopus offers the broadest and most consistent coverage of the governance-oriented, behavioural, and institutional research streams that define the epistemological scope of this review. This is a deliberate epistemological boundary, not a default convenience choice; future reviews should extend the corpus by incorporating Web of Science and IEEE Xplore to capture cross-disciplinary perspectives, as noted in Section 7. Accordingly, studies published predominantly in computer science and engineering venues (e.g., IEEE Transactions, ACM Digital Library) fall outside the epistemological scope of this review. This exclusion is scope-driven and discipline-specific rather than a quality-based judgement: technical venues have produced foundational and increasingly governance-aware XAI contributions (including work that explicitly addresses fairness and institutional accountability in financial modelling but their disciplinary conventions prioritise computational benchmarking as the primary evaluative criterion, which is distinct from the governance, behavioural adoption, and institutional outcome constructs examined here [70,71]. This review is therefore intentionally positioned as a complement to the algorithmic literature, not a replacement of it [70,71]. The open-access filter was applied for two distinct and deliberate reasons. First, epistemological transparency: open-access publications afford full-text verifiability by all readers and reviewers, which is directly consistent with the reproducibility ethos underlying PRISMA 2020 and the SPAR-4-SLR protocol adopted in this review. Second, empirical representativeness: open-access publication in business, management, and FinTech governance has grown substantially since 2020 [68,69], meaning that the open-access corpus now captures a high and growing proportion of high-quality governance-oriented XAI research. Nonetheless, this filter constitutes a formal limitation in that subscription-access publications in the field remain excluded; this is acknowledged and future reviews are encouraged to replicate this study's search strategy without the open-access restriction.

Also, this review is a qualitative SLR/meta-framework study; formal quantitative risk-of-bias tools (e.g., ROBIS) were not applied. Moreover, since this review employs qualitative thematic synthesis rather than quantitative meta-analysis, no formal effect measures were calculated.

The final counts and sequential filtering of studies are documented in the PRISMA Flow Diagram (Section 2.4), and the retained corpus's analytical structuring followed the SPAR-4-SLR protocol, as described in Section 2.5.

2.4. Study Selection Procedure (PRISMA 2020)

The study screening protocol adhered to PRISMA 2020 reporting standards [65]. Figure 3 shows the PRISMA Flow Diagram of the sequential stages of the identification, screening, and eligibility and inclusion of featured articles.

The initial Scopus retrieval yielded 12,600 records. After the implementation of predefined filters for language, document type, publication status, disciplinary scope, and Open Access option, 12,344 records were excluded from the database through an automated refinement process, resulting in 256 articles.

Following a quality assessment against the Australian Business Deans Council (ABDC) Journal Quality List, only articles published in journals listed on this list proceeded to the

next step in the process. Once this purification step was completed, 99 articles met the established relevance and quality thresholds and comprised the final literature corpus for the review.

Following the approach adopted in comparable framework-based SLRs, the initial title and abstract screening was conducted by the primary researcher. To mitigate the risk of selection bias inherent in single-reviewer screening, a systematic double-check of excluded records was performed against the inclusion and exclusion criteria detailed in Section 2.3, and all borderline cases were retained for full-text review rather than excluded at the screening stage. Formal inter-rater reliability testing (e.g., Cohen's Kappa) was not undertaken, given the structured and protocol-driven nature of the coding process. Coding consistency was ensured through the use of a predefined coding framework applied uniformly across all studies, supported by an initial pilot coding exercise and subsequent iterative validation to maintain consistency across ADO and TCCM classifications. The complete PRISMA 2020 Checklist was complied with, and is also supplied with the manuscript as Supplementary Material.

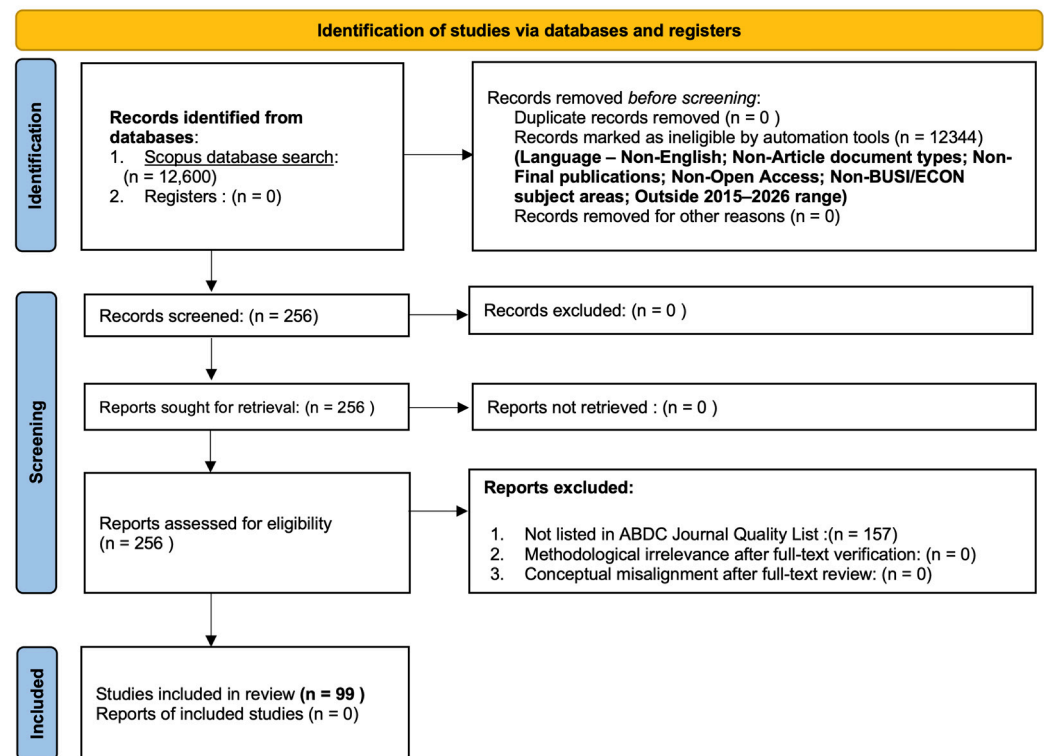


Figure 3. PRISMA Flow Diagram (2020) (adapted from PRISMA 2020 framework [65]).

2.5. SPAR-4-SLR Protocol

The review's analytical structure is based on the SPAR-4-SLR framework [66]. The standardised stages of the process, viz. Assemble, Arrange, and Assess are shown in Figure 4 (SPAR-4-SLR Framework).

The use of the SPAR-4-SLR SLR approach complements the transparency ensured by PRISMA by systematically categorising retained studies by their conceptual relationships, thereby developing a coherent conceptual organisation. The combination of methodology and organisation creates greater methodological clarity and a deeper level of interpretation for the synthesis produced [72,73].

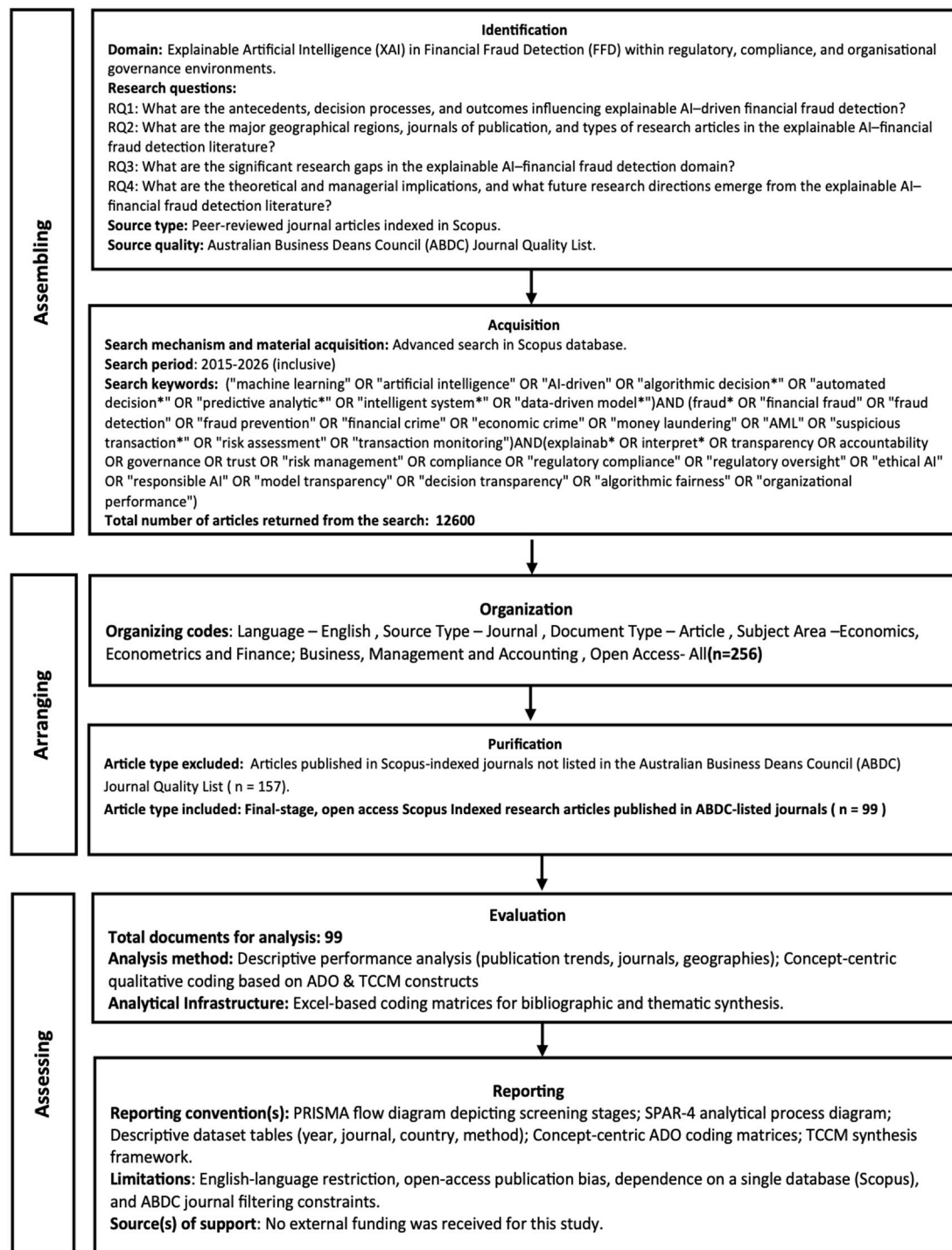


Figure 4. SPAR-4-SLR Framework (adapted from [66]).

3. FinTech-Oriented XAI System Architecture

A FinTech-oriented architecture for explainable AI-driven fraud detection systems is outlined here, extending the conceptual analysis through four integrated layers. The architecture integrates four core layers:

- (i) Data Layer: High-volume transactional, behavioural, and alternative data streams;
- (ii) Model Layer: Machine learning and deep learning algorithms (e.g., GNNs, ensemble models);

- (iii) Explainability Layer: Post hoc and intrinsic XAI methods (e.g., SHAP, LIME, counterfactual explanations);
- (iv) Decision and Governance Layer: Human-in-the-loop validation, regulatory reporting, and audit compliance mechanisms.

The four-layer architecture depicted in Figure 5 reflects how explainability operates as an embedded governance component within FinTech infrastructures rather than as an external add-on. It also highlights the feedback loop among model outputs, human oversight, and regulatory requirements, ensuring continuous model validation and alignment with governance (vide Figure 5). Subsequently, Table 2 differentiates different XAI Techniques based on type, strength, limitation & use in fintech.

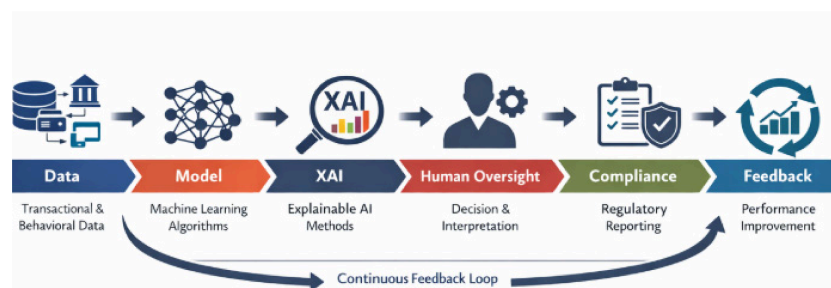


Figure 5. XAI-Enabled FinTech Fraud Detection Architecture (Authors' Compilation).

Table 2. Comparison of XAI Techniques.

Method	Type	Strength	Limitation	Use in FinTech
SHAP	Post hoc (global and local)	Consistent, theoretically grounded	Computationally expensive	Credit scoring, fraud explanation
LIME	Post hoc (local)	Model-agnostic, simple	Instability across runs	Transaction-level fraud analysis
Decision Trees	Intrinsic	Highly interpretable	Lower accuracy vs complex models	Rule-based fraud detection
Rule-based Models	Intrinsic	Transparent and auditable	Limited scalability	AML compliance systems
Counterfactual Explanations	Post hoc	Actionable insights	Complex to generate	Loan approval explanations

4. Analysis and Synthesis

4.1. Annual Scientific Production

The descriptive and thematic findings are presented below, structured sequentially by publication trends, theoretical foundations, methodological approaches, and sectoral applications. Figure 6 charts annual output from one publication in 2015 to twenty-five in 2025, illustrating rapid growth in scholarly interest in AI-driven financial systems. Post 2020, the notable inflexion point observed may be attributed to the implementation of XAI by financial institutions for financial modelling and compliance analytics [61,65]. These developments coincide with the introduction of regulatory schemes (e.g., the GDPR and the EU AI Act) that spurred increased academic interest in developing transparent, auditable, and ethically compliant financial algorithms [20,21].

Key Insight: The most important pattern in the production data is the inflexion point after 2020. It corroborates that academic interest in XAI governance is not organically technology-driven, but institutionally catalysed—directly coincident with the implementation timeline of the GDPR and the EU AI Act. This temporal alignment provides the basis

for the governance-first framing of the current review, while also defining the epistemological setting of the ADO analysis in Section 4.

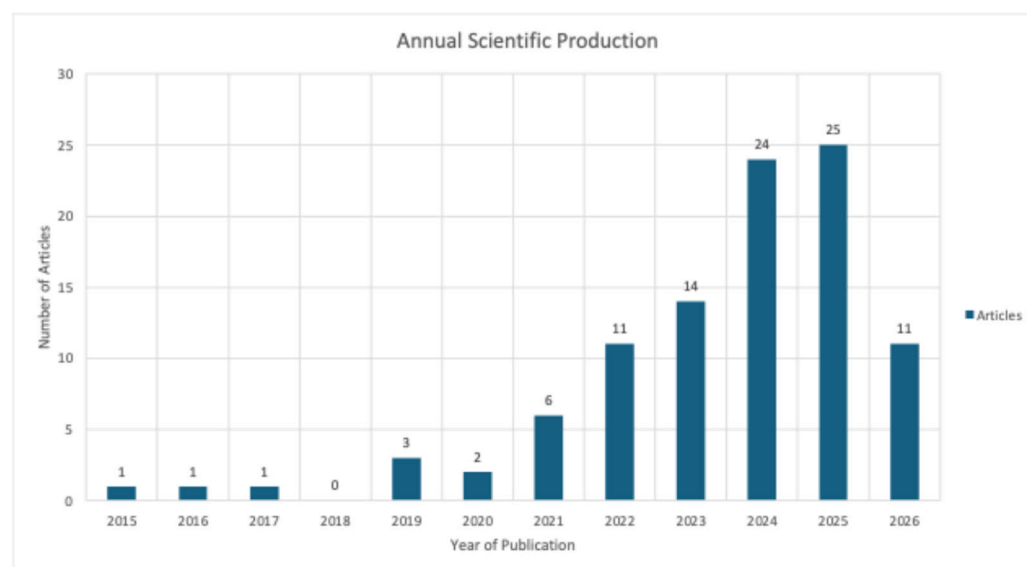


Figure 6. Annual Scientific Production (2015–2026).

4.2. Source-Journal Distribution and Indexing Profile

Table 3 documents 99 articles across 70 distinct journals. A and A-ranked journals account for almost 48% of the corpus, concentrated in *Decision Support Systems* (A*), *Technological Forecasting and Social Change* (A), and the *Journal of Forecasting* (A), confirming the field’s consolidation within high-quality decision science and financial analytics arenas [74,75].

Key Insight: The distribution of 70 journals in 99 articles in the corpus indicates the absence of a dominant publication outlet for research on XAI-in-finance governance. This fragmentation is not a weakness—it reflects the genuinely interdisciplinary character of the field—but it also explains why no previous review has been able to synthesise the literature from within any single disciplinary tradition. It offers a direct methodological justification for the multi-theory ADO framework used in this paper.

Table 3. Journal-Wise Distribution of Included Studies and ABDC Ranking Classification.

Journal	Indexing and ABDC Ranking	No. of Articles	% of Dataset
<i>Decision Support Systems</i>	Scopus, A*	6	6.06%
<i>Technological Forecasting and Social Change</i>	Scopus, A	4	4.04%
<i>Journal of Forecasting</i>	Scopus, A	4	4.04%
<i>Journal of Banking Regulation</i>	Scopus, C	3	3.03%
<i>International Journal of Production Research</i>	Scopus, A	3	3.03%
<i>Finance Research Letters</i>	Scopus, A	3	3.03%
<i>Journal of Money Laundering Control</i>	Scopus, C	3	3.03%
<i>Technology in Society</i>	Scopus, C	2	2.02%
<i>Technovation</i>	Scopus, A	2	2.02%
<i>Journal of Economic Surveys</i>	Scopus, A	2	2.02%
<i>Information Resources Management Journal</i>	Scopus, C	2	2.02%
<i>Journal of Organisational and End User Computing</i>	Scopus, B	2	2.02%
<i>International Journal of Production Economics</i>	Scopus, A	2	2.02%
<i>Journal of Behavioural and Experimental Finance</i>	Scopus, A	2	2.02%
<i>International Review of Finance</i>	Scopus, A	2	2.02%
<i>Economies</i>	Scopus, C	2	2.02%

Table 3. Cont.

Journal	Indexing and ABDC Ranking	No. of Articles	% of Dataset
<i>Energy Economics</i>	Scopus, A*	2	2.02%
<i>International Journal of Information Systems and Supply Chain Management</i>	Scopus, C	2	2.02%
<i>Electronic Commerce Research and Applications</i>	Scopus, B	2	2.02%
<i>Engineering, Construction and Architectural Management</i>	Scopus, A	2	2.02%
<i>IEEE Transactions on Engineering Management</i>	Scopus, A	2	2.02%
<i>Electronic Markets</i>	Scopus, A	2	2.02%
<i>International Transactions in Operational Research</i>	Scopus, B	2	2.02%
<i>Computational Management Science</i>	Scopus, B	1	1.01%
<i>Sustainable Futures</i>	Scopus, C	1	1.01%
<i>Systems Research and Behavioural Science</i>	Scopus, A	1	1.01%
<i>Journal of Industrial Relations</i>	Scopus, A	1	1.01%
<i>Transportation Research Part A: Policy and Practice</i>	Scopus, A*	1	1.01%
<i>Journal of Applied Economic Sciences</i>	Scopus, C	1	1.01%
<i>Journal of Construction Engineering and Management</i>	Scopus, A*	1	1.01%
<i>Risk Management and Insurance Review</i>	Scopus, C	1	1.01%
<i>Emerging Markets Review</i>	Scopus, A	1	1.01%
<i>Administrative Sciences</i>	Scopus, C	1	1.01%
<i>International Journal of Business Information Systems</i>	Scopus, C	1	1.01%
<i>South African Journal of Business Management</i>	Scopus, C	1	1.01%
<i>Business Strategy and the Environment</i>	Scopus, A	1	1.01%
<i>Journal of Industrial and Management Optimisation</i>	Scopus, B	1	1.01%
<i>Journal of Risk and Insurance</i>	Scopus, A	1	1.01%
<i>International Journal of Economics and Financial Issues</i>	Scopus, C	1	1.01%
<i>Current Issues in Auditing</i>	Scopus, B	1	1.01%
<i>European Journal of Health Economics</i>	Scopus, A	1	1.01%
<i>Transportation Research Part E: Logistics and Transportation Review</i>	Scopus, A*	1	1.01%
<i>Managerial and Decision Economics</i>	Scopus, B	1	1.01%
<i>Management Review Quarterly</i>	Scopus, B	1	1.01%
<i>Marine Policy</i>	Scopus, A	1	1.01%
<i>Borsa Istanbul Review</i>	Scopus, B	1	1.01%
<i>Strategic Management Journal</i>	Scopus, A*	1	1.01%
<i>Production Planning and Control</i>	Scopus, A	1	1.01%
<i>Financial Innovation</i>	Scopus, B	1	1.01%
<i>Journal of Business Economics</i>	Scopus, B	1	1.01%
<i>Accounting Research Journal</i>	Scopus, B	1	1.01%
<i>Indian Journal of Marketing</i>	Scopus, C	1	1.01%
<i>Economic Modelling</i>	Scopus, A	1	1.01%
<i>Public Administration Review</i>	Scopus, A	1	1.01%
<i>Journal of Cases on Information Technology</i>	Scopus, C	1	1.01%
<i>Management Science</i>	Scopus, A*	1	1.01%
<i>Journal of Econometrics</i>	Scopus, A*	1	1.01%
<i>Journal of the Operational Research Society</i>	Scopus, A	1	1.01%
<i>Journal of Global Information Management</i>	Scopus, A	1	1.01%
<i>European Business Organisation Law Review</i>	Scopus, B	1	1.01%
<i>Oxford Review of Economic Policy</i>	Scopus, A	1	1.01%
<i>Journal of Advanced Transportation</i>	Scopus, B	1	1.01%
<i>Tourism Management</i>	Scopus, A*	1	1.01%
<i>International Journal of Project Management</i>	Scopus, A*	1	1.01%
Total		99	100.00%

4.3. Geographical Research Distribution

As summarised in Table 4, research activity is globally dispersed yet regionally concentrated. China (56 publications), the United Kingdom (45), and Australia (28) lead contributions, followed by the United States (26) and Germany (22). Together, European economies account for over 40 per cent of the total sample, underscoring the regulatory and institutional impetus driving AI-finance research [76]. Additionally, the noticeable increase in scholarly publications from India (12), Malaysia, and Turkey after 2022 shows that emerging markets are beginning to participate more in integrating AI into their capital markets [4,8].

Key Insight: The geographical concentration of output in China, the UK and Australia—three jurisdictions with distinct but equally stringent AI regulatory environments—affirms that XAI governance research is institutionally rather than technically driven. Regulatory pressure is the most plausible common factor across otherwise diverse research cultures, and reinforces the normative antecedent cluster identified in Section 5.2.

Table 4. Country-Wise Distribution of Scholarly Contributions in the Final Sample.

Country	Number of Articles	Country	Number of Articles	Country	Number of Articles
China	56	Romania	6	Iran	2
UK	45	Ukraine	6	Ireland	2
Australia	28	Belgium	5	Malaysia	2
USA	26	Thailand	5	Serbia	2
Germany	22	Brazil	4	South Korea	2
Spain	19	South Africa	4	Austria	1
Italy	18	Sweden	4	Cyprus	1
India	12	Turkey	4	Morocco	1
France	10	Canada	3	Pakistan	1
Finland	7	Qatar	3	Singapore	1
Netherlands	7	Colombia	2	Slovenia	1
Switzerland	7	Greece	2	United Arab Emirates	1

4.4. Theoretical Foundations

As illustrated in Table 5, the analysed corpus revealed 15 unique theoretical frameworks. Amongst the 15 frameworks, the Resource-Based View (RBV) and Institutional Theory are the two dominant frameworks (13 studies using the RBV and 12 studies using Institutional Theory), reflecting a dual emphasis on capability exploitation and institutional compliance [29,46]. Other frameworks that contribute to understanding human-AI dyads, in addition to human–AI adaptive decision architectures, include Decision Theory (11 studies) and Socio-Technical Systems Theory (8 studies) [77,78]. In addition, the new and emerging frameworks of Algorithmic Fairness, Risk Management, and Trust in Automation continue to advance the paradigm toward ethical governance and operational transparency [79,80].

Key Insight: RBV (13 studies) and Institutional Theory (12 studies) prevalence over TPB (4 studies) in the reviewed corpus points to an underlying gap: the field theorises XAI as a resource or legitimacy mechanism and not yet a behaviourally governed adoption decision. The present study precisely fills this gap by positioning TPB as the main theoretical lens in the ADO meta-framework.

Table 5. Theoretical and conceptual frameworks underpinning the included studies and their frequency distribution.

Sr. NO	Theory/Framework	References	No. of Studies
1	Resource-Based View (RBV)	[8,10,36,46–48,71,74,81–85]	13
2	Institutional Theory	[4,7,9,20,21,29,49–51,86–88]	12
3	Decision Theory and DSS	[6,37,40,52–54,77,78,89–91]	11
4	Agency Theory	[18,20,26,55,92–95]	9
5	Socio-Technical Systems Theory	[7,8,12,18,77,96–98]	8
6	Technology Acceptance (TAM/UTAUT)	[37,38,47,99–102]	7
7	Signal Theory/Info Asymmetry	[103–109]	7
8	Fuzzy Set Theory and Logic	[81,83,95,110,111]	5
9	Algorithmic Fairness Theory	[11,60,79,112,113]	5
10	Risk Management/PIDE Theory	[24,80,85,114]	4
11	Theory of Planned Behaviour (TPB)	[47,50,59,115]	4
12	Fraud Diamond/Triangle Theory	[12,59,92,116]	4
13	Reliability Theory (XAI Stability)	[23,28,43]	3
14	Innovation Resistance Theory (IRT)	[37–39]	3
15	Modern Portfolio Theory (MPT)	[80]	1

4.5. Methodological Orientation

As shown in Table 6, the reviewed studies exhibit a methodologically balanced architecture. The majority of studies utilised quantitative designs ($\approx 60\%$), which included econometric modelling, algorithm experimentation, and simulation-based validation [22,117]. Fifteen of these studies utilised hybrid or mixed methods ($\approx 15\%$), combining statistical performance assessments with expert judgement to provide dual validation of technical and cognitive reliability [52,53]. Additionally, 12% of studies utilised qualitative or case-based approaches to provide contextual information regarding governance, regulation, and organisational trust following the paradigms established by Turkesen et al. (2024 [7]; Hickman & Petrin, 2021 [29]. This three-pronged methodological triad—quantitative precision, interpretive grounding, and practical validation—establishes a sound empirical framework for the development of explainable AI for fraud detection and decision intelligence within the finance industry.

Key Insight: The methodological imbalance is compounded by the quantitative preponderance ($\sim 60\%$) and the almost non-existence of longitudinal governance studies (methodological imbalance outlined in the Introduction). The field under review provides algorithmic evidence but little institutional evidence, a gap that the present review directly addresses through qualitative thematic synthesis.

Table 6. Methodological Classification and Research Design: Distribution of the Included Studies.

Sr. NO	Type of Study	No. of Studies	Methodological Sub-Types	References	
1	Quantitative	Quantitative: Algorithmic/Experimental	41	Deep learning, GBDT, ensemble methods, and simulation-based stress testing.	[5,11,12,22–24,26,28,34–36,42,45,59,81,83–85,95,96,104,106–109,112,114,116–126]
2		Quantitative: Survey/Econometric	18	Structural equation modelling (SEM), GMM, and cross-sectional regression.	[10,17,37,38,46,47,50,79,92,93,99,103,105,115,127–130]
3	Qualitative	Qualitative: Case Study and Interviews	12	Multiple-case analysis, expert elicitation, and semi-structured interviews.	[3,4,7,9,29,39,49,60,86,88,97,100]

Table 6. Cont.

Sr. NO	Type of Study	No. of Studies	Methodological Sub-Types	References
4	Hybrid/Mixed-Methods	15	Combining human-expert benchmarking with AI performance scores (blind peer review).	[6,8,25,35,38,52,53,80,89,110,131–135]
5		13	Systematic literature reviews (SLR) and PRISMA-based framework development.	[8,18,20,21,25,51,55,71,74,80,82,133,136]

4.6. Sectoral Distribution of Empirical Application

The industrial mapping, as shown in Table 7, shows a concentration of research across seven sectors. Banking and Credit (37 studies) is the most prevalent sector of study due to the development of models for credit scoring, loan default prediction, and SME risk optimisation, all of which rely on interpretable XAI [28,95]. E-commerce and Payments (12 studies) focus on fast-moving, graph-driven, fraud-triage environments [23,119]. RegTech and AML (12 studies) focus on automating compliance and identifying suspicious activities as specified by Directive (AMLD 6) [4,7]. The new sectors of study include Investment and Portfolio (nine studies), Insurance (four studies), ESG Corporate Governance (seven studies), and Supply Chain Risk (six studies), highlighting how AI has transitioned from simply supporting transaction oversight, facilitating institutional governance [9,104]. The asymmetries seen across sectors indicate that the field is developing two trajectories: precision in prediction in finance and accountability in governance interpretation.

Key Insight: The most policy-relevant finding to emerge from the sectoral data is the asymmetry between Banking and Credit (37 studies) and fairness-adjacent sectors such as Insurance (4 studies) and ESG (7 studies). This asymmetry suggests that XAI governance research is concentrated where financial risk is most measurable but underinvested precisely where algorithmic bias carries the greatest societal consequence—a tension directly addressed by the Ethical Governance Protocol mechanisms in Section 5.2.2.

Table 7. Sectoral Distribution of Included Studies, Associated Decision Contexts, and Technical Capability Focus.

Sr. NO	Industry Sector	No. of Studies	Decision Context	Technical Capability Focus	References
1	Banking and Credit	37	Credit scoring, loan default, and SME creditworthiness assessment.	Risk Sensitivity: Balancing predictive accuracy with semantic interpretability for auditors.	[8,17,20–22,24,25,28,34,35,40,45,48,51,54,55,58,74,77–79,81,82,90,91,94,95,98,102,106,107,109,117,123,125,136,137]
2	E-commerce and Payments	12	Real-time transaction fraud, fake review detection, and user behavioural embedding.	High-Velocity Triage: Processing high-frequency graphs and clickstreams in real-time.	[23,36,42,43,83–85,102,108,112,119,138,139]
3	RegTech and AML	12	AML forensics, suspicious activity reporting (SAR), and compliance monitoring.	Regulatory Auditability: Mapping Article-level law into traceable control criteria.	[4,5,7,12,29,86–88,96,99,100]
4	Investment and Portfolio	9	Asset allocation, crypto market risk, and pension fund optimisation.	Uncertainty Quantification: Modelling tail-risk spillovers and scenario trees.	[80,103,109,114,126,140–143]

Table 7. *Cont.*

Sr. NO	Industry Sector	No. of Studies	Decision Context	Technical Capability Focus	References
5	Insurance and Claims	4	Underwriting fraud, high-cardinality claim management, and premium misrepresentation.	Complexity Handling: Managing high-dimensional categorical data and label-free detection.	[11,113,118,122]
6	Corporate and ESG	7	ESG disclosure accountability, CEO deception, and takeover market risk.	Linguistic Forensics: Using NLP/Sentiment Analysis to detect obfuscation and salience.	[9,46,50,93,104,105,116]
7	Supply Chain and Logistics	6	SCF payment risk, chemical supply chain fraud, and logistics disruptions.	Interdependence Modelling: Mapping nonlinear evolutionary trajectories across stakeholders.	[6,125,129,132–134]

4.7. Data Typology and Complexity Landscape

Table 8 depicts four primary categories of data used in studies. Structured financial datasets are included in 44 studies. These datasets are based on high-volume, structured financial data, such as stock market indices, transactional ledgers, and financial statements, which are high-dimensional and exhibit nonlinear dependencies [116,125]. Behavioural or perceptual data on organisational trust in decision-making are reported as survey feedback or Delphi panel responses in 22 studies [38,47]. There are 18 studies on unstructured alternative dataset types (text, sentiment analysis, and satellite log) that require the implementation of real-time natural language processing (NLP) technologies and graph embeddings [9,104]. Thus, the addition of the legal/systemic corpus (15 studies), comprising legislation/regulatory texts and associated policies, enables the synthesis and interpretation of the dataset as a whole [4,20].

Key Insight: The presence of structured financial data (44 studies) and behavioural/perceptual data (22 studies) in the same corpus validates that XAI in finance is a dual technical and human problem. This dual data landscape gives an empirical basis for STS Theory integration in the STR model—both subsystems, technical and social, are empirically populated.

Table 8. Data Typology, Source Characteristics, and Complexity Profile of the Included Studies.

Sr. NO	Data Category	No. of Studies	Data Source and Type	Data Velocity and Complexity	References
1	Secondary: Structured Financial	44	Stock market indices, bank transaction logs, credit bureau reports, and IFRS/GAAP statements.	High Complexity: Multi-dimensional numerical data requiring feature engineering for non-linearity.	[5,11,22,24,26,28,35,36,42,45,58,81,83,84,95,96,106–109,114,116–126,141]
2	Primary: Behavioural and Perceptual	22	Semi-structured interviews, Likert-scale surveys, and expert Delphi panels.	Low Velocity: Static snapshots of human intent, trust levels, and organisational resistance.	[3,7,10,17,37,38,46,47,49,50,79,92,93,99,100,103,105,115,127–130]
3	Alternative: Unstructured and Synthetic	18	Sentiment from earnings calls, social media text, satellite sensor logs, and GAN-generated fraud samples.	Ultra-High Velocity: Requires real-time NLP or Graph embedding; high noise-to-signal ratio.	[6,12,23,39,55,60,85,86,89,96,104,111,112,132,134,137,144]

Table 8. *Cont.*

Sr. NO	Data Category	No. of Studies	Data Source and Type	Data Velocity and Complexity	References
4	Secondary: Legal and Systematic	15	Regulatory texts (GDPR/EU AI Act), previously published literature, and policy White Papers.	High Complexity: Qualitative density; requires thematic synthesis and cross-reference mapping.	[4,8,9,20,21,25,29,51,74,80,82,88,97,133,136]

4.8. Decision-Type Taxonomy

Decision contexts are distributed across six primary archetypes as shown in Table 9. Risk-Predictive (31 studies) and Operational-Triage (22 studies) dominate, validating XAI's functional core in financial-loss mitigation through semi-autonomous decision support [19,20]. Diagnostic-Forensic studies (16%) emphasise post-event traceability, linking AI outputs to legal reasoning [5,104]. Strategic-Governance applications (12%) extend interpretability into ESG accountability and trustworthy AI policy [9,29]. Prescriptive-Tactical and Fairness-Adjustive segments (circa 10% and 6%, respectively) remain under-represented, highlighting the need for bias correction and decision optimisation [79,113]. The typology confirms a research orientation anchored in predictive validity, yet gradually expanding toward fairness and strategic governance—both critical to the ADO meta-framework.

These findings highlight the increasing convergence between explainable artificial intelligence and FinTech innovation. The concentration of studies in banking, payments, and RegTech domains indicates that explainability is becoming a foundational requirement for scalable digital financial services, particularly in environments characterised by high transaction volumes, regulatory scrutiny, and real-time decision-making.

Key Insight: The dominance of Risk-Predictive (31 studies) and Operational-Triage (22 studies) decision types, compared to the underrepresentation of Fairness-Adjustive (6 studies) and Prescriptive-Tactical (9 studies) types, reveals a field that has operationalised explainability as a precision instrument but not yet fully institutionalised it as a fairness and strategic governance tool. The principal finding of the descriptive synthesis is that, across the dominant decision types in the corpus viz. Risk-Predictive and Operational-Triage contexts; XAI adoption has achieved technical maturity while governance institutionalisation remains uneven and sector-dependent, most evidently advanced in Banking and Credit and RegTech, but substantively underdeveloped in Fairness-Adjustive and Strategic-Governance applications. This finding directly motivates the ADO framework analysis in Section 4 and the STR model introduced in Section 6.

Table 9. Decision Typology and associated Autonomy Levels across the included studies.

Sr. NO	Decision Type	No. of Studies	Decision Context	Decision Autonomy Level	References
1	Risk-Predictive	31	Forecasting probability of default (PD), credit distress, and crash risks.	Semi-Autonomous: AI generates risk scores; humans set the institutional risk thresholds.	[8,22,24,26,28,58,80,81,95,106–109,117,121,124–126,133,141]
2	Operational-Triage	22	Real-time flagging of fraudulent transactions, clickstreams, and AML suspicious activities.	High-Autonomous: Automated blocking/triaging with human forensic override.	[5,7,12,23,35,36,42,43,59,88,96,118,119,122,144]

Table 9. *Cont.*

Sr. NO	Decision Type	No. of Studies	Decision Context	Decision Autonomy Level	References
3	Diagnostic-Forensic	16	Root-cause analysis of fraud signatures, deceptive reviews, and CEO earnings call obfuscation.	Human-Centric: AI identifies salient patterns; humans provide the legal/contextual rationale.	[3,5,21,39,49,50,60,84,86,92,93,97,100,104,105,116]
4	Strategic-Governance	15	Corporate ESG accountability, public sector financial auditing, and Trustworthy AI policy.	Executive-Level: AI insights inform long-term institutional policy and risk posture.	[4,9,18,20,21,25,29,49,51,55,74,82]
5	Prescriptive-Tactical	9	Optimising staff planning, ambulance relocation, and cash management buffer adjustments.	Dynamic Support: AI prescribes specific mitigation tactics; humans approve/execute.	[6,85,89,114,129,132,134,135,145]
6	Fairness-Adjustive	6	Correcting bias in subprime lending and mis-calibration in alternative behavioural data.	Corrective: AI acts as a monitor to recalibrate human underwriting rules for equity.	[10,11,59,60,79,123]

5. Content Analysis Using the ADO Framework

This study integrates the current literature regarding Explainable Artificial Intelligence (XAI) for financial fraud detection using the Antecedents–Decisions–Outcomes (ADO) framework to categorise the relevant academic research into various categories based on the antecedents of influence used in decision-making processes that lead to the resulting outcomes at the organisational or societal levels. The model has been used as a reliable and effective means of developing integrated systematic reviews of management, information systems and financial technology research, thereby producing coherent synthesis and supporting future research directions [143,144,146,147] that can be applied in this case through the classification of technological, behavioural, and institutional elements that together create the definition of explainability as a strategic governance capability within the financial ecosystem [148,149].

Consistent with the theoretical hierarchy established in Section 1, the resultant conceptual framework is behaviourally governed by the **Theory of Planned Behaviour (TPB)** as its primary lens across all three ADO layers, with RBV, Institutional Theory, and Decision Theory enriching the antecedent, outcome, and decision layers, respectively.

5.1. The Theory of Planned Behaviour (TPB)

Ajzen's Theory of Planned Behaviour provides a foundation for understanding behavioural patterns driven by cognitive, normative, and control mechanisms in relation to the implementation and adoption of explainable artificial intelligence (XAI) systems in financial decision-making [150–152]. The theory postulates that behaviour is guided by three psychological determinants: attitude toward the behaviour, subjective norms, and perceived behavioural control, which collectively shape behavioural intention and subsequent action [150]. The framework has been validated across a wide range of technological and managerial contexts, including e-governance, cloud computing, and digital banking [153,154], demonstrating its applicability to AI-driven decision-making environments.

This study uses attitudinal antecedents as perceptions of the interpretability, transparency, and perceived usefulness of XAI systems; subjective norms refer to the various

regulatory, professional, and ethical pressures that enforce explainability obligations; and perceived behavioural control encompasses data-governance readiness, technological competence, and organisational support [150–152,155]. These three antecedent constructs provide a basis for how financial managers, auditors, and regulators view the ability of an XAI system to serve as a reliable, legitimate source, given its current governance.

As previously mentioned, the decision-making aspect of TPB involves turning those beliefs into actionable steps, such as developing an algorithmic validation system and creating interpretability audits or compliance reports [7,23,106]. These behaviours are examples of intention-based organisational behaviour, driven by both cognitive assessments and regulatory enforcement, in which human logic ties with machine logic under institutional supervision [40,53].

Outcomes from these behaviours include trust calibration, regulatory compliance, operational reliability, and ethical accountability, all of which result from the social and organisational value created by XAI adoption [9,29,39,79]. The impact of XAI on organisations supports TPB's premise that attitudes, norms, and perceived behavioural control lead to a long-lasting form of trustworthiness at the varying levels of complexity found within the financial industry.

The ADO Meta Framework (Figure 7) represents a new epistemic approach that connects psychological factors in the Theory of Planned Behaviour [148–150] to the ADO Model's causal chain, which illustrates how attitudes translate into behaviour in decision-making and ultimately produce institutional outcomes with respect to accountability and trust. Explainability is framed within the Behaviourally Governed ADO Meta Framework as a behavioural mechanism that provides stability to the human–AI interface and legitimises the use of automated financial audit systems when combined with algorithmic governance [2,43,153].

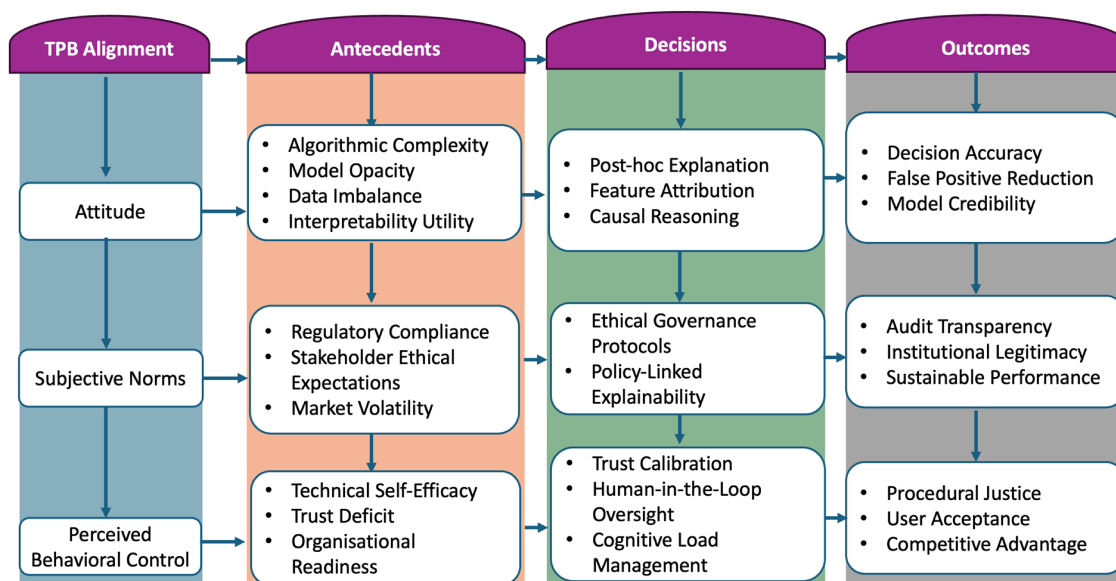


Figure 7. Behaviourally Governed ADO Meta-Framework for Explainable AI in Financial Fraud Detection.

5.2. Antecedents

As summarised in Table 10, this section deconstructs the **Antecedents (A)** driving Explainable AI (XAI) adoption in Financial Fraud Detection (FFD). Grounded in the **Theory of Planned Behaviour (TPB)**, we categorise these drivers into three pillars: **Attitude**, **Subjective Norms**, and **Perceived Behavioural Control (PBC)**. This mapping reveals how internal evaluations, external mandates, and organisational capacities converge to form the strategic intention to implement interpretable machine learning.

Table 10. Antecedents of the Explainable AI-Driven Financial Fraud Detection Model.

Theory of Planned Behaviour (TPB) Dimension	Antecedent Construct	Conceptual Description and Empirical Findings (<i>N</i> = 99)	Representative Financial Fraud Detection Contexts	References
Attitude	Algorithmic Complexity	The use of high-dimensional architectures (DNN, GNN, XGBoost) creates nonlinear logic that prevents direct human tracing.	Transactional fraud alerts	[22,23,35,36,117]
	Model Opacity	The inherent ‘black-box’ nature of neural logic erodes the user’s ability to justify interventions, creating an epistemic gap.	Credit risk and loan audits	[17,18,20,32,156,157]
	Data Imbalance	Highly skewed fraud datasets (1:1000 ratio) lead to biased boundaries, requiring XAI to validate minority class logic.	SME loan default prediction	[43,45,104,158]
	Interpretability Utility	Perceived value of XAI in converting statistical probabilities into actionable intelligence for forensic investigators.	Risk assessment strategy	[8,34,58,107,117]
Subjective Norms	Regulatory Compliance	Global mandates (GDPR, EU AI Act) transition XAI from a technical elective to a mandatory legal “Right to Explanation.”	AML forensics and SAR	[4,7,12,21,29]
	Ethical Expectations	Pressure from stakeholders to ensure fraud models do not introduce systemic bias against protected demographic groups.	ESG and fair lending	[2,9,10,46,79]
	Market Volatility	Dynamic shifts in fraud patterns demand a sound rationale to ensure models have not drifted or failed during market stress.	Capital market analytics	[90,91,159,160]
Perceived Behavioural Control	Technical Self-Efficacy	The organisation’s internal literacy and specialised skillsets are required to deploy and maintain auditable XAI pipelines.	AI literacy and HR capital	[37,38,56,161,162]
	Trust Deficit	Managerial scepticism toward automated results limits adoption unless the AI provides context-rich justifications.	Bank wire forensics	[40,90,91,102,163]
	Organisational Readiness	The firm’s digital maturity, learning culture, and available infrastructure support computationally intensive XAI tools.	FinTech vs Legacy Banks	[56,57,162]

5.2.1. Attitude Toward XAI Adoption

The attitudinal dimension represents the internal evaluative beliefs of financial organisations regarding the utility, risks, and necessity of XAI integration.

- **Algorithmic Complexity**—Expertise in designing and running highly conventional analytics, such as linear statistical models, has enabled the transition to using more cognitively complex, higher-dimensional architectures for predictive modelling. Examples include Deep Neural Networks (DNN), Graph Neural Networks (GNN), and Extreme Gradient Boosting (XGBoost). The increased complexity of these methodologies has exponentially increased the ability to predict potential fraud patterns [23,36]. Nonlinear logic is one of the structural impediments to tracing fraud alerts back to a specific person. [22]. The research indicates that as modelling complexity increases, forensic investigators have greater difficulty verifying hidden-layer interactions manually, suggesting that additional effort is necessary to leverage XAI to reconcile the model's mathematical representations with their corresponding variables in the financial domain [35,117,121].
- **Model Opacity**—Due to their black box nature, AI algorithms create an epistemological distance between the output produced by the algorithm and the rationale/business case for the manager's decision [32,164]. For example, in credit risk assessments and loan audits, this opacity makes it difficult for loan auditors to explain the basis for AI-generated discrepancies to internal review panels and regulatory bodies, leading to negative responses within an organisation to the use of automated systems. [20,137]. For these reasons, it is argued that XAI is a necessary corrective antecedent that enables logical validation of interventions when making decisions about high-stakes financial transactions [45,157].
- **Data Imbalance**—Many organisations are implementing XAI for various reasons, one of which is the imbalance present in fraud datasets. For instance, there might only be 1 fraud occurrence for every 1000 legitimate transactions [104]. The resulting heavily unbalanced distribution, therefore, leads to incorrect judgments of the relationships between the different classes of observations [158]. Organisations use XAI to determine whether their fraud detection notifications are based on evidence of actual fraud or are merely the result of statistical distributions driven by data imbalance [112]. If XAI provides local explanation(s), then investigators can ensure that any bias in the model's "learning" did not occur due to over-sampling or noise [43,45].
- **Interpretability Utility**—According to Liu et al. (2024) [58], XAI's value lies in its ability to transform abstract statistical probabilities into actionable business intelligence. As researchers view interpretability more as a practical utility than a high-end feature, they believe it can improve the speed and accuracy of triaging fraud alerts based on risk [117]. Recent research indicates that once XAI identifies the 'causal drivers' of a fraud alert, the perceived value of the AI system will substantially increase, thereby promoting the organisation's favourable intention to scale its use [8,34,109].

5.2.2. Subjective Norms in the XAI Ecosystem

Subjective norms represent the external pressures and perceived social/legal mandates that compel financial institutions to adopt explainable frameworks, described as follows:

- **Regulatory Compliance**—Due to global regulatory mandates like GDPR (Article 22) and the European Union's AI Act, 'explainable artificial intelligence' has moved from being a discretionary choice in the development of an AI system to a 'losing' legal need [4,29]. The 'right to explanation' has established a normative landscape requiring financial institutions to justify their automated decisions in Anti-Money Laundering (AML), forensic accounting, and suspicious activity report (SAR) filing. [12]. Failure to

maintain a logic trail that can be audited introduces significant legal liability risk and drives the normatively significant adoption of XAI within organisations for regulatory compliance [7,21,61].

- **Ethical Expectations**—Firms are pressured by different stakeholders, including investors, civic advocacy groups, and others, to build fraud detection models that do not create systemic bias against certain protected groups [46]. In the context of fair lending and ESG, XAI is employed to meet these normative ethical expectations by providing evidence that racial and gender-related variables were not used as proxy inputs into the fraud prediction process [2,9]. This value alignment supports an organisation's legitimacy as an institution and its continued operation [10,79].
- **Market Volatility**—Due to the changing nature of fraud schemes, there is a need for rationale to validate that models have not become obsolete or failed [159]. As defined by market norms, to avoid serious liquidity problems or customer losses caused by the cascading effect of false positives, trust-by-design is required in high-volatility scenarios [117]. Organisations have begun using XAI to provide rationales that can substantiate model performance when traditional statistical benchmarks are unstable [90,160].

5.2.3. Perceived Behavioural Control (PBC)

PBC reflects the internal organisational enablers—resources, skills, and infrastructure—that determine the perceived ease or difficulty of implementing XAI, discussed as follows:

- **Technical Self-Efficacy**—According to an article by Kim & Kim (2026) [37], successful implementation depends on an organisation's internal AI literacy and the expertise of data science employees. A firm with higher levels of technical self-efficacy than other firms is more successful in implementing complex interpretability pipeline techniques, such as SHAP or GNN Explainers, without significantly affecting its existing fraud detection processes [38]. Organisations determine how much control they have over recalibrating AI logic based on their capabilities [56,161,162].
- **Trust Deficit**—The persistent scepticism among upper management regarding automated 'black-box' alerts represents a major obstacle to perceived control [102]. This lack of trust creates a behavioural bottleneck: if (upper) managers do not feel they have control over how the machine has made its 'decision', they will be less likely to act on the output [90,163]. XAI builds a bridge to restore that behavioural control by providing contextually rich justifications that align the machine's output with the manager's domain knowledge [40,91].
- **Organisational Readiness**—A firm's digital infrastructure and governance culture maturity level drives its ability to host computationally intensive XAI tools [56]. A ready organisation has the necessary hardware for processing complex explanations of the data driving those tools and a "learning culture" that prepares analysts to utilise the rationales for XAI each day in their triage work [162]. Research indicates that most FinTech firms are more ready to use XAI than traditional banks, as their back-end infrastructure is built with auditable-by-default data pipelines [57,98].

5.3. Decisions

These mechanisms operationalise the strategic intention to adopt XAI by bridging the gap between high-dimensional machine logic and organisational forensic requirements. As shown in Table 11, following the Theory of Planned Behaviour (TPB), they are categorised by the specific behavioural antecedents they address.

Table 11. Decision Processes within the ADO Model of Explainable AI in Financial Fraud Detection.

Theoretical Alignment		Key Antecedent	Mechanism/Mediator	Functional Role in Decision Making	Illustrative Contexts	References
Attitude	Utility	Algorithmic Complexity; Model Opacity	Post-hoc Explanation	Translating deep math into local importance scores (SHAP/LIME) for individual transactional alert triage.	Card Fraud Triage	[23,28,43,165]
	Utility	Data Imbalance; Interpretability Utility	Feature Attribution	Global weight mapping of variables (e.g., transaction velocity) to validate the core model logic.	Credit/HR Analytics	[117,121,166–168]
	Logic	Data Imbalance	Causal Reasoning	Generating counterfactual “what-if” scenarios to clarify the causal path behind automated denials or flags.	Loan Recourse and Appeals	[11,45,60,123,169]
Subjective Norms	Social	Ethical Expectations	Ethical Governance Protocols	Embedding bias-monitoring and demographic parity metrics directly into the real-time auditing lifecycle.	Corporate Audits	[11,27,61,79,112]
	Legal	Regulatory Compliance	Policy-Linked Explainability	Aligning internal AI reporting standards with international regulatory benchmarks (e.g., Basel III/IV).	SAR/AML Compliance	[7,21,61,164,170]
Perceived Behavioural Control	Control	Trust Deficit	Trust Calibration	The iterative process of adjusting managerial reliance on AI based on the veracity of the provided logic.	High-stakes Trading	[52,53,90,102,163]
	Control	Organisational Readiness	Human-in-the-Loop (HITL)	Establishing forensic oversight where experts validate AI flags before final disposition.	AML Investigation	[40,77,78,89,98]
	Ease of Use	Technical Self-Efficacy; Organisational Readiness	Cognitive Load Management	Designing XAI interfaces that present complex data in digestible, user-centric visualisations.	Internal Bank Audits	[40,57,171]

5.3.1. Addressing Technological Antecedents (Attitude and Utility)

The primary decision mechanisms used to alleviate Algorithmic Complexity and Model Opacity are through the translation of “deep math” into human-understandable local and global explanations.

- **Post hoc Explanation (SHAP/LIME)**—Localised feature importance scores are used to provide real-time validation of high-stakes alerts by identifying key features within each transactional alert triage. Key features, such as high transaction velocity and geographic anomalies, can support analysts by providing real-time insight into the validity of alerts from transactionally detected or planned activity [23,28,43].
- **Feature Attribution**—By mapping the global weights of all variables across the model to determine overall data imbalances and maximise the value of interpretation and use, companies can reduce the risk of evaluating non-legitimate data artefacts. This is particularly important in the context of credit and HR analytics. To ensure that the model is not based solely on invalid variables, companies work to capture legitimate predictors that reinforce the overall value and usefulness of the entire model [117,166–168].
- **Causal Reasoning**—Causal reasoning is a method of “what if” counterfactual scenario analysis that can help investigators determine the reasons for an automated denial of a loan [11]. In these cases, investigators must show a causal relationship to justify an automated denial; they use counterfactual reasoning to determine the minimal change required to alter a person’s loan application outcome [45,60].

5.3.2. Addressing Environmental Antecedents (Subjective Norms and Legal Compliance)

Institutional mechanisms ensure that the AI ecosystem aligns with external pressures, including ethical standards and strict regulatory mandates.

- **Ethical Governance Protocols**—Organisations incorporate metrics for bias monitoring and demographic parity into their real-time auditing processes in response to social pressures and ethical expectations [79]. This eliminates discrimination against protected classes arising from the logic used to detect fraud and maintains institutional integrity [11,27,112].
- **Policy-Linked Explainability**—Organisations use policy-linked explainability to align their internal AI reporting requirements with international regulatory standards, such as Basel III/IV, and to comply with legal and regulatory requirements [61]. The formalisation of these standards ensures that the SAR (Suspicious Activity Report) and AML (Anti-Money Laundering) processes of all organisations are based on traceable and auditable evidence [7,21,164,170].

5.3.3. Addressing Organisational Antecedents (Perceived Behavioural Control and Readiness)

Cognitive and socio-technical mechanisms are deployed to enhance the firm’s internal capabilities and address psychological barriers to adoption.

- **Trust Calibration**—To mitigate the issues caused by a Trust Deficit within an organisation, developing and consistently implementing an iterative Trust Calibration process through explainable AI (XAI) is a means of allowing managers to enhance or reduce their dependence upon AI, using the truthfulness and clarity of the logical rationale provided by the AI. By providing the “reliability” associated with an explanation, organisations will reduce automation bias [53,90,163].
- **Human-in-the-Loop (HITL) Oversight**—Organisations with high levels of Readiness will implement HITL oversight, thereby providing a means to retain Control by the human subject matter expert. With AI serving as a high-speed filtering mechanism, expert forensic investigators will add the final “interpretive layer” to confirm the flags

before any can be disposed of. This requires algorithms to operate at high speed while still taking into account the intuitive judgement of an expert in their field [40,77,78,98].

- Cognitive Load Management—To enhance Ease of Use, organisations are developing user-centric visualisations that aid investigators in processing complex information from XAI systems, enabling quicker interpretation and decision-making [171]. As a result, investigators are more likely to experience higher levels of Technical Self-Efficacy. They are not negatively affected by the amount of data presented to them while conducting internal bank audit tasks under high pressure [40,57].

5.4. Outcome

As shown in Table 12, the outcomes are categorised based on their alignment with the Theory of Planned Behaviour (TPB), demonstrating how specific decision mechanisms fulfil attitudinal, normative, and control-related goals.

Table 12. Outcomes of Explainable AI–Enabled Financial Fraud Detection within the ADO Framework.

Theoretical Alignment		Decision Construct	Outcome Construct	Description of Impact and Empirical Results	Predominant Domains	References
Attitude	Precision	Interpretative Post hoc Explanation Feature Attribution Causal Reasoning	Decision Accuracy	Improved precision by identifying and eliminating spurious or ‘hallucinated’ data correlations.	Banking and Stock Markets	[172–174]
	Efficiency		False Positive Reduction	Significant improvement in triage efficiency, reducing resource waste in manual reviews.	Fraud Management	[85,111,114,117,175]
	Reliability		Model Credibility	Enhanced technical reliability through verifiable and reproducible logic paths.	Credit and FinTech	[2,32,121,176]
Subjective Norms	Compliance	Institutional Ethical Governance Protocols Policy-Linked Explainability	Audit Transparency	Traceable, defensible justifications for AI decisions provided to regulatory authorities.	AI Auditing	[4,29,177,178]
	Legitimacy		Institutional Legitimacy	Strengthened social licence to operate through demonstrated adherence to ESG standards.	ESG and Public Policy	[9,46,61,179]
	Sustainability		Sustainable Performance	Long-term organisational resilience through ongoing adherence to compliance frameworks.	Corporate Governance	[30,81,180,181]
Perceived Behavioural Control	Fairness	Cognitive Trust Calibration Human-in-the-Loop (HITL) Cognitive Load Management	Procedural Justice	Heightened perception of fairness and trust when automated flags are rationalised.	HR Management and Credit	[27,78,79,181]
	Usage		User Acceptance	Reduction in “innovation anxiety” and increased willingness to depend on AI tools.	Retail and Public Admin	[40,57,171]
	Strategy		Competitive Advantage	Transforming explainability from a compliance burden into a dynamic asset for innovation.	FinTech Innovation	[10,31,55,162]

5.4.1. Outcomes of Interpretative Mechanisms (Attitude and Precision)

Technical and operational outcomes validate the Attitude that XAI is a functional necessity for precision-driven fraud detection.

- **Decision Accuracy**—Interpretative mechanisms improve decision accuracy by detecting and deleting false or fictitious data relationships, thereby increasing model accuracy. This is especially important in Banking and the Stock Market, where determining how to justify the cause of a “red flag” before acting on it is critical for financial integrity [23,36,172,182].
- **False Positive Reduction**—By leveraging XAI rationale to assess alerts raised by investigators rapidly, XAI enables substantial improvements in operational efficiency by preventing the waste of forensic resources, allowing Fraud Management teams to focus on investigating only the highest-probability fraudulent activities [85,111,114,117,175].
- **Model Credibility**—Explainability fosters reliability by providing verifiable and reproducible logic paths. In Credit and FinTech, the model’s technical credibility assures that it will produce high-quality results, not only through statistical analysis but also through a forensic perspective [176,183–185].

5.4.2. Outcomes of Institutional Mechanisms (Subjective Norms and Compliance)

Institutional outcomes satisfy Subjective Norms by aligning the organisation with external legal and social mandates.

- **Audit Transparency**—Provides an auditable trail that supports a verifiable, defensible rationale, as required by regulatory bodies for compliance [4]. Verification is essential for AI auditing, as black-box outputs do not provide a legally sufficient rationale for Compliance SARs or AML reporting [29,177,178].
- **Institutional Legitimacy**—By demonstrating compliance with environmental, social, and governance (ESG) parameters and the ethical standards of AI governance, companies establish legitimacy. Demonstrating fraud detection logic as unbiased and traceable through verifiable processes provides a social licence to operate, confirming that they are operating legitimately [9,46,170].
- **Sustainable Performance**—As long as organisations demonstrate ongoing compliance with regulatory framework(s), they remain sustainable and provide long-term organisational resilience. With respect to corporate governance, alignment with regulatory standards prevents legal risks and ensures the ongoing continuity of operations [30,81,180].

5.4.3. Outcomes of Cognitive Mechanisms (Perceived Behavioural Control and Fairness)

Behavioural outcomes reflect the restoration of control and the psychological acceptance of AI as a decision-making partner.

- **Procedural Justice**—Automated alerts accompanied by a clear, logical rationale allow users/employees to perceive greater fairness [78]. The importance of perceived Fairness is significant in both HR Management and Credit, as transparency in decision-making helps reduce negative perceptions of algorithmic bias [27,79].
- **User Acceptance**—Increased transparency decreases “innovation anxiety” and thus increases usage. Users in Retail and Public Administration will have more confidence in using AI-assisted tools if they believe they can exercise cognitive control over their outputs [40,57,171].
- **Competitive Advantage**—Organisations that view explainability as a compliance obligation rather than as a dynamic strategy potentially gain a significant market advantage [55]. Companies engaged in FinTech Innovation, capable of quickly de-

playing auditable, trusted AI, can achieve a distinct competitive advantage in the marketplace [10,31,162].

6. Discussion

The foundational theoretical problem this study addresses is the absence of a behaviourally governed, institutionally embedded model that explains how and why organisations adopt explainability as a governance capability, rather than as a technical feature, under conditions of simultaneous algorithmic complexity, regulatory pressure, and organisational uncertainty. Prior frameworks for XAI governance, including the NIST AI Risk Management Framework (NIST AI RMF), the EU AI Act's Trustworthy AI principles, and the Explainability Maturity Model proposed by [21], address this challenge from a compliance or performance-benchmarking perspective but do not model the behavioural, recursive, and socio-technical mechanisms through which explainability becomes institutionalised. The STR (Stability–Transparency–Reliability) model introduced in this study is a direct theoretical answer to that gap: it is behaviourally grounded in TPB, recursive through the DCV-based learning loop, and institutionally embedded through the Triadic STS Alignment model. The following discussion interprets the empirical findings through this theoretical lens, explicitly distinguishing descriptive patterns from theoretical implications.

The present study investigates four research questions regarding how Explainable Artificial Intelligence (XAI) can be applied to detect financial fraud through a behaviourally anchored ADO framework combined with the Theory of Planned Behaviour (TPB). The discussion interprets the findings in relation to the research questions, highlighting how explainable AI transitions from a technical feature to a governance imperative within financial systems. Through a systematic review of 99 articles, this study seeks to offer insights into the antecedents, decision processes, and outcomes that potentially drive the evolution of explainability from a technical utility toward a fundamental governance necessity. Additionally, this review identifies trends in publications, research types, and themes that are currently underinvestigated. The findings can provide a basis for a clearer, more refined theoretical framework and managerial guidance for XAI-enabled financial systems.

6.1. Antecedents of the Explainable-AI Financial-Decision Model

Addressing RQ1, the findings reveal three distinct antecedent dimensions—Attitude, Subjective Norms, and Perceived Behavioural Control—derived from the TPB. Each captures the underlying drivers influencing the adoption and operationalisation of explainability in financial decision contexts.

6.1.1. Attitudinal Antecedents

The attitudinal antecedent cluster extends the Resource-Based View by demonstrating that XAI literacy and infrastructure is a forensic governance capability—not just a technical resource—with direct implications for how financial institutions prioritise AI governance investment. In the data, this theoretical claim is operationalised through four constructs: algorithmic complexity, model opacity, data imbalance, and interpretability utility. Within the dataset, four constructs dominate: algorithmic complexity, model opacity, data imbalance, and interpretability utility. Complex nonlinear algorithms such as DNNs, GNNs, and XGBoost increase epistemic uncertainty [22,23,36]. Additionally, neural networks have limitations that limit their justification for use in credit-risk audits. [18,32]. The existing imbalance between classes in fraud data increases the risk of bias and requires post hoc interpretability to validate predictions for the minority class [43,104]. Collectively, these antecedents reflect managerial awareness that interpretability is improving the forensic analysis of algorithmic prediction outputs [34,112]. In theory, this cluster progresses

RBV: the discovery that companies with greater AI literacy and XAI infrastructure are in a stronger position to leverage interpretability as a governance asset expands the RBV concept of dynamic capability from general organisational resources into the specifically forensic realm of algorithmic accountability.

6.1.2. Subjective-Norm Antecedents

The normative antecedent cluster confirms Institutional Theory's core prediction: under conditions of regulatory and stakeholder pressure, organisations adopt explainability not as a technical elective but as a legitimacy-securing behaviour. Three empirical constructs substantiate this theoretical position: regulatory compliance mandates (EU AI Act, GDPR Article 22), stakeholder ethical expectations, and market volatility. The EU AI Act and Article 22 of the GDPR constitute a formalised "right to explanation," which is included in the regulatory requirements [4]. In addition, stakeholder ethical expectations put even greater pressure on model designers to ensure demographic fairness [9,79]. Finally, several studies confirm that market volatility forces institutions to verify that model drift does not impair interpretive stability under stress [90,117]. The theoretical implication is significant: the normative antecedent cluster shows that the adoption of explainability is not primarily a technical choice, but rather a legitimacy-seeking behaviour that is directly in line with Institutional Theory's prediction that organisations conform to institutional pressures in order to obtain a social licence to operate.

6.1.3. Perceived Behavioural Control Antecedents

The Perceived Behavioural Control antecedents extend TPB's construct of control beyond resource availability into what this study terms Interpretative Agency, defined as the cognitive capacity to forensically validate automated logic rather than merely operate the tool. Two manifestations of this extended construct are empirically supported: capability confidence (technical self-efficacy) and infrastructure maturity (organisational readiness). Adoption can only be successful if specialised analytical skills are available, along with an overall learning culture that can sustain auditable AI pipelines [37,38,56]. In contrast, there is ongoing scepticism about opaque automation unless the rationale for using AI is contextually rich and verifiable [40,102]. Therefore, two manifestations of control beliefs, viz., capability confidence and infrastructure maturity, jointly shape behavioural intention to implement explainability [57,98]. This builds on the original idea of Perceived Behavioural Control by the Theory of Planned Behaviour (TPB) to extend its scope from resource availability to what this study terms Interpretative Agency, the cognitive capacity to forensically validate automated logic, not simply to access or operate the tool.

6.2. Decision Aspect of the ADO Model

The decision mechanisms identified in the corpus confirm Decision Theory's prediction that interpretability tools are adopted when they restore the human decision-maker's sense of causal control over otherwise opaque systems—not primarily because they improve model performance. Three mechanism categories operationalise this theoretical claim across the three TPB antecedent dimensions.

The mapping of attitudinal antecedents to post hoc explanation mechanisms (SHAP, LIME, counterfactual reasoning) confirms Decision Theory's prediction that decision-makers will adopt interpretability tools when they perceive these tools as reducing epistemic uncertainty in high-stakes choices [11,24,39,56]. The theoretical advance here is that explainability is not adopted because it improves model performance—it is adopted because it restores the human decision-maker's sense of causal control over an otherwise opaque system.

Within the subjective-norm antecedent, the organisational institutionalisation of ethical-governance protocols establishes ethical and policy-based explainability. These frameworks include embedding fairness and demographic-parity checks into compliance workflows to withstand regulatory scrutiny [7,61].

Perceived behavioural control focuses on the development of trust in human-in-the-loop systems. Hybrid models of human auditors validating automated alerts before disposition assist with both accountability and agility [40,53]. Cognitive-load-management interfaces further help individuals comprehend complex AI logic by providing user reports or dashboards that reduce “explanation fatigue” due to complexity and/or by providing a means of comprehension [53,170,185].

6.3. Outcomes of the ADO Model

The outcome findings validate the “Justified Utility” extension of Decision Theory proposed in Table 12: in high-stakes financial decisions, utility encompasses not only the predictive outcome but the traceability and verifiability of the causal path leading to it. Three outcome clusters provide the empirical basis for this theoretical claim. They are precision outcomes, compliance outcomes, and fairness outcomes, with each supporting a different level of behaviour along the TPB model.

Precision outcomes (reduced false positives, improved decision accuracy, enhanced model credibility) confirm that XAI delivers measurable operational returns beyond regulatory compliance [23,172]. The three precision outcomes documented in Table 12, reduced false positives, improved decision accuracy, and enhanced model credibility, collectively support the “Justified Utility” extension of Decision Theory formalised in Section 7.1.3: the evidence that XAI-enabled institutions achieve superior triage efficiency precisely because they can trace and verify the causal path of each prediction confirms that utility in high-stakes financial decisions is inseparable from explainability.

Outcomes related to the subject norms group include the following: increased transparency in audits; increased legitimacy of the institution; and alignment of the institution with sustainability goals. Systems that can explain how they arrived at their answers will provide more defensible responses from a regulatory perspective and provide a clearer picture of compliance with established ethical and ESG guidelines [4,9].

Outcomes related to perceived behavioural control include perceptions of procedural fairness, user acceptance, and perceptions of competitive advantage. Decisions perceived as rational will increase user trust, reduce anxiety about innovation, and turn the concept of explainability from a compliance cost into a competitive advantage [55,162,182].

6.4. Geographical Distribution, Journals, and Research Typologies

To answer the second research question (RQ2), bibliometric mapping indicates that research into detecting financial fraud using XAI technology is primarily occurring in developed economies (North America, Europe, and East Asia) but expanding into emerging markets, where institutions are also beginning to study this area. Most articles are published in leading journals across the domains of Information Systems, Decision Sciences, and Financial Technology, including but not limited to the *European Journal of Operational Research*, *Journal of Business Ethics*, and *Journal of Financial Regulation and Compliance*.

The most common methodology is quantitative (algorithmic and econometric), with approximately 60% using this method. The next most common are mixed-methods and theoretical papers focusing on governance and ethics; however, there has been a recent trend to integrate interdisciplinary approaches between behavioural governance/finance and data science. These findings collectively address RQ2, confirming the field’s methodological evolution and geographical expansion.

From the point of view of a digital FinTech ecosystem—digital banking platforms, payment systems and RegTech solutions—the findings indicate that explainability is not a feature but a fundamental requirement for scalable digital financial systems.

6.5. *Tensions, Contradictions, and the Novelty of the STR Model*

Three unresolved tensions emerge from the reviewed corpus that existing governance frameworks have not adequately addressed, and which the STR model is specifically designed to resolve.

6.5.1. Tension 1: Accuracy Versus Interpretability:

The most persistent tension in the literature is the trade-off between predictive accuracy and interpretability [21,70,71]. High-performance models (DNNs, GNNs) achieve fraud detection accuracy exceeding 90% but produce explanations of limited fidelity. Existing frameworks such as the NIST AI RMF and the EU AI Act's Trustworthy AI guidelines acknowledge this trade-off but do not resolve it; they require transparency without specifying how transparency can be achieved without sacrificing performance. The STR model's Stability component directly addresses this: by treating explanation consistency across market cycles as a governance metric independent of model accuracy, the STR model reframes the trade-off. An organisation can maintain high-accuracy models while demonstrating governance-level stability through SHAP consistency benchmarking, converting a technical dilemma into a manageable institutional process.

6.5.2. Tension 2: Regulatory Mandate Versus Operational Feasibility:

The EU AI Act's 'right to explanation' standard creates a legal obligation that is technically challenging in millisecond-latency fraud detection environments, where generating post hoc explanations may introduce processing delays incompatible with real-time triage [4,17,57]. This tension is absent from [21] Explainability Maturity Model, which evaluates explainability as a static organisational capability rather than a real-time operational constraint. The STR model's Transparency component addresses this by distinguishing between transactional-level transparency (SHAP-based, produced post hoc) and audit-level transparency (Model Card-based, produced at the SAR stage), enabling compliance at the appropriate temporal granularity without compromising operational speed.

6.5.3. Tension 3: Static Compliance Versus Recursive Governance:

All three comparator frameworks—NIST AI RMF, EU AI Act, and [21] treat XAI governance as a point-in-time compliance state: an organisation either meets the standard or it does not. None of them model the recursive dynamic through which XAI outcomes improve future organisational readiness. The STR model's Reliability component, grounded in the DCV-based recursive learning loop, is the only framework among the four that explicitly positions XAI as a self-reinforcing governance catalyst. The STR model's primary novelty claim is precisely this conversion: governance transformed from a static audit condition into a dynamic, iterative learning process: from compliance-as-destination to governance-as-trajectory.

Two methodological limitations qualify the scope of these theoretical claims and warrant explicit acknowledgement at the level of the findings rather than as a standalone section. First, while corpus-level quality was ensured through the ABDC Journal Quality List filter, risk of bias was not formally assessed at the individual-study level using a quantitative instrument; the absence of ROBIS or an equivalent tool means that study-level methodological weaknesses may not be fully accounted for in the thematic synthesis, and readers should interpret the ADO categorisations with this in mind. Second, as a systematic literature review dependent on the quality and scope of the underlying studies,

the findings may not fully capture emerging real-time applications of XAI in rapidly evolving FinTech environments, a limitation that the Recursive Frontier research agenda Section 8 is specifically designed to address.

7. Implications

7.1. Theoretical Implications

Addressing RQ4, the theoretical and managerial implications of this review emerge directly from the three STR model extensions and the ADO synthesis, with each implication traceable to the specific empirical findings documented in Sections 4 and 5. As summarised in Table 13, three theoretical advances emerge from the reviewed corpus, each extending an established framework into the domain of algorithmic governance of the present study by transitioning the academic discourse from a simplified “Black-Box” discussion toward a structured understanding of Socio-Technical Recalibration. The core theoretical contribution of this research is the introduction of the STR (Stability–Transparency–Reliability) Model. This model serves as an integrative framework that advances the literature on technology acceptance and algorithmic governance across three primary subject areas.

7.1.1. Advancing the Theory of Planned Behaviour: The STR Component of Reliability

A primary theoretical objective of this research is to offer a possible refinement to the understanding of Perceived Behavioural Control (PBC), specifically in the context of high-stakes algorithmic environments. This paper presents the STR Model to facilitate “Interpretative Agency,” a critical component of PBC. The findings indicate that just having access to AI tools will not allow for meaningful, effective organisational actions; rather, the ability for a user to forensically validate the rationale of machines through both Trust Calibration and Human-in-the-Loop (HITL) processes demonstrates that PBC is based on the user’s ability to verify the rationale of the machine [40,102,183]. This will help transition the Theory of Planned Behaviour (TPB) from a “Tech Usage” paradigm to one based on “Cognitive Governance”, in which users can sustain their use of an AI tool if they have an adequate understanding of why it works as intended [90,163].

7.1.2. The STR Model and Triadic Alignment in Socio-Technical Systems (STS)

The second named extension, Triadic STS Alignment, introduced in Contribution 3 of Section 1, carries a direct implication for how financial institutions should design their governance architectures: the traditional two-subsystem STS model (Social and Technical) is structurally insufficient for regulated AI environments and must be formally extended to incorporate a third, non-negotiable pillar—the Regulatory and Ethical Subsystem. While traditional STS focuses on the co-optimisation of Social (people) and Technical (tools) subsystems, our research indicates that these subsystems struggle to achieve equilibrium in regulated industries without the formal bridge provided by Transparency and Stability. We propose the “Triadic STS Alignment” model, in which XAI serves as the unifying factor, aligning algorithmic performance with institutional accountability and societal expectations [4,29,61]. This triadic extension implies that financial institutions cannot achieve genuine XAI governance compliance by optimising social and technical factors alone—the regulatory–ethical pillar must be institutionally embedded as a co-equal subsystem, not treated as an external constraint.

7.1.3. Formalising the Recursive Learning Loop via Interpretative Agility

Lastly, this study seeks to provide a grounded evolution of the Dynamic Capabilities View (DCV) by formalising the Recursive Feedback Loop within the ADO framework. The STR Model conceptualises XAI as a “Learning Catalyst” that assists in creating “Interpre-

tative Agility,” with XAI seen as an active process rather than a static output. Whereas Interpretative Agency operates at the level of actor cognition and validation capacity, Interpretative Agility operates at the level of organisational adaptation and learning capability. The study identifies that building Institutional Trust (Outcome) can shift the baseline for Organisational Readiness (Antecedent) in future cycles. The recursive relationship further indicates that XAI can be regarded as a potential self-reinforcing strategic asset that advances innovation capabilities by improving institutional resilience over the long term and reducing barriers to future AI innovation.

Table 13. Proposed Theoretical Advancements and Refinements.

Theoretical Foundation	Traditional Tenet	STR Model Refinement/Advancement	Proposed Mechanism (ADO-Linked)	Reference
Theory of Planned Behaviour (TPB)	PBC is a measure of resource availability and ease of use.	Reliability (Interpretative Agency): Redefines control as the cognitive ability to forensically validate automated logic.	Trust Calibration and Cognitive Load Management	[40,90,102,163]
Socio-Technical Systems (STS)	Optimisation depends on both the Social and Technical subsystems.	Transparency (Triadic Alignment): Proposes a “Regulatory/Ethical” pillar as a non-negotiable stabiliser.	Policy-Linked Explainability and Ethical Protocols	[4,29,61,79]
Dynamic Capabilities (DCV)	Firms integrate and reconfigure internal/external competences.	Stability (Recursive Learning): Positions XAI as a catalyst reconfiguring future adoption readiness.	Recursive Feedback Loop	[31,55–57]
Institutional Theory	Organisations achieve legitimacy through “Box-Ticking” mimicry.	Strategic Legitimacy: Transforms passive compliance into an active differentiator for innovation.	Audit Transparency and Governance Protocols	[10,17,61,162]
Decision Theory	Decisions maximise expected utility.	Justified Utility: Proposes that utility includes both the outcome and the traceability of the path.	Post hoc Explanation and Causal Reasoning	[11,23,36]

7.2. Managerial and Policy Implications

The three theoretical extensions formalised in Section 6.1, namely Interpretative Agency, Triadic STS Alignment, and the Recursive Learning Loop (operationalised through Interpretative Agility), translate directly into managerial imperatives, shifting practice from passively deploying AI to actively governing it within a socio-technical environment. From a FinTech perspective, these implications are particularly relevant for digital banking platforms, payment systems, and RegTech solutions, where automated decision-making must be both efficient and transparent. The integration of explainability into these systems supports not only compliance but also customer trust and market competitiveness. As presented in Table 14, CROs and operations managers may need to focus on Policy-Linked Explainability as the primary means of connecting black-box outputs to their legal obligations under the GDPR and the EU AI Act [4,21,61]. To reduce the Trust Deficit that prevents successful AI integration, senior management can make HITL oversight and Trust Calibration part of their operating procedures and allow forensic experts their

Interpretative Agency when following up on automated alerts [77,102,163]. The use of Ethical Governance Protocols can also support compliance while providing a means to establish Institutional Legitimacy and to satisfy stakeholder expectations regarding algorithmic fairness and demographic parity [9,46,79]. Finally, by maximising Cognitive Load Management and Technical Self-Efficacy through focused training, organisations might be able to turn explainability into a Sustainable Competitive Advantage, leading to higher False Positive Reduction rates and greater operational resilience during volatile market conditions [31,55,57,111].

Table 14. Managerial and Policy Implications Blueprint.

Focus Area	Implication Type	Primary Actors/Stakeholders	Strategic Recommendations	Research-Driven Insights	Practice-Based Implementation Ideas
Algorithmic Accountability	Policy	Regulatory Bodies (GDPR, EU AI Act), CROs	Shift from “Passive Compliance” to “Accountability-by-Design” to meet the legal Right to Explanation.	Regulatory mandates are transitioning XAI from a technical elective to a legal requisite [4,21].	Develop standardised “Model Cards” that provide post hoc SHAP rationales for every SAR submission.
Operational Precision	Managerial	Data Scientists, Fraud Analysts	Utilise Interpretative Mechanisms to eliminate spurious correlations in highly imbalanced fraud datasets.	XAI identifies “hallucinated” correlations that lead to costly false positives [23,36].	Integrate SHAP/LIME dashboards directly into the real-time transaction triage workflow.
Socio-Technical Trust	Managerial	Chief Information Officers (CIOs), Investigators	Institutionalise Human-in-the-Loop (HITL) oversight to calibrate human reliance and prevent automation bias.	Trust is not binary; it requires constant calibration through transparent logic [102,163].	Implement “Co-Piloting” sessions where analysts forensically validate AI flags using counterfactual reasoning.
Ethical Governance	Policy and Social	ESG Committees, Ethics Boards	Embed Ethical Protocols to monitor demographic parity and mitigate systemic bias in automated lending.	Stakeholder expectations demand that fraud models remain free from proxy-variable discrimination [79].	Perform “Fairness Audits” quarterly using XAI to identify disparate impacts on protected groups.
Resource Efficiency	Managerial	Operations Managers, Forensic Leads	Leverage XAI-driven triage to optimise the allocation of expensive forensic expertise.	XAI significantly reduces the time spent on manual review of low-probability alerts [111,117].	Create a tiered alert system: “Automated Clear,” “XAI-Explained Triage,” and “Complex Forensic Review.”
Innovation Strategy	Managerial	Chief Strategy Officers, FinTech Leads	Transform explainability from a “Compliance Burden” into a “Dynamic Strategic Asset” for market differentiation.	Firms with auditable AI achieve higher institutional trust and faster product approvals [31,55].	Market “Transparent AI” as a core brand value to enhance user acceptance and customer loyalty.
Cognitive Readiness	Policy	HR Directors, Learning and Development	Prioritise XAI Literacy training to bridge the gap between technical self-efficacy and managerial control.	Organisational readiness is the ultimate bottleneck for effective XAI implementation [37,40].	Launch an “AI-Forensics Certification” for staff to master interpretability tools and cognitive load management.

8. Future Research Directions

Addressing RQ3, the research gaps identified through the ADO synthesis and the bibliometric mapping point to five under-investigated frontiers, as systematised in Table 15: technical real-time XAI scalability, the legal admissibility of AI-generated explanations, the fairness–accuracy–interpretability trilemma, optimal trust calibration thresholds, and longitudinal validation of the STR model’s recursive dynamics. This section outlines areas for future academic investigation, as revealed by gaps identified in the extant literature. The present review’s focus on the business and management literature, while theoretically coherent, leaves open the productive question of how technically oriented XAI research in computer science venues converges with or diverges from the governance findings reported here. Cross-database SLRs integrating IEEE, ACM, and arXiv corpora alongside Scopus would allow future scholars to map the full technical–institutional spectrum of XAI in fraud detection. Further, as shown in Table 15, the research agenda proposed by this study calls for a shift in focus to the longitudinal, real-time, and humanistic nature of the current paradigm in social scientific research, away from the interpretation of static models of human behaviour.

Table 15. Proposed Future Research Directions.

Research Frontier	Identified Gap	Proposed Research Questions	Suggested Methodological Approach	Target Theoretical Contribution	References
Technical Frontier	Real-time XAI for high-velocity transaction streams.	How can local explanation fidelity be maintained in millisecond-latency streaming architectures without degrading detection speed?	Experimental Design; Simulation of Streaming Architectures	Technical Utility and Efficiency	[23,36]
Technical Frontier	Standardised forensic-grade fidelity metrics.	What objective metrics best quantify the “ground truth” alignment between a post hoc explanation and the underlying non-linear logic?	Quantitative Benchmarking; Mathematical Validation	Model Credibility and Reliability	[35,165,184]
Institutional Frontier	Legal admissibility and evidential weight of XAI.	To what extent do XAI rationales (e.g., SHAP plots) meet the “Daubert Standard” or equivalent legal criteria for evidence in SAR disputes?	Legal Hermeneutics: Qualitative Case Study Analysis	Regulatory Compliance and Audit Transparency	[4,12,61]
Institutional Frontier	The Fairness-Accuracy-Interpretability “Trilemma.”	How does the enforcement of demographic parity protocols specifically alter the causal pathways revealed by counterfactual reasoning?	Multi-Objective Optimisation; Algorithmic Auditing	Ethical Norms and Procedural Justice	[27,61,79]
Behavioural Frontier	Optimal trust calibration vs automation over-reliance.	What is the threshold of information density where XAI shifts from empowering an analyst to causing cognitive “over-reliance”?	Lab Experiments; Eye-Tracking; Behavioural Triage Tasks	PBC and User Acceptance	[40,90,102,163]
Behavioural Frontier	Cognitive load and UX design for auditors.	Which visualisation archetypes most effectively minimise information fatigue for forensic investigators during high-pressure triage?	User Experience (UX) Research; Longitudinal Field Studies	Technical Self-Efficacy and Ease of Use	[57,171]
Strategic Frontier	Long-term ROI and competitive advantage.	Does the adoption of auditable AI lead to a quantifiable reduction in fraud-loss ratios and an increase in institutional market value?	Longitudinal Financial Analysis; Event Study Methodology	Dynamic Capabilities (DCV) and Strategy	[10,31,55]
Recursive Frontier	Verification of the Organisational Learning Loop.	How does the successful realisation of XAI outcomes over time reconfigure a firm’s future readiness for disruptive AI innovations?	Qualitative In-Depth Interviews; Longitudinal Surveys	Recursive Learning and Organisational Readiness	[56,57,162]

9. Conclusions

This study sought to advance the governance of XAI in financial fraud detection by synthesising a 99-article corpus through a TPB-anchored ADO–TCCM meta-framework. The findings suggest that explainability is no longer a discretionary technical feature but a behaviourally governed, institutionally mandated strategic imperative—shaped by attitudinal, normative, and control-based antecedents and operationalised through post hoc explanation tools, ethical governance protocols, and human-in-the-loop oversight. The core theoretical contribution, the STR (Stability–Transparency–Reliability) Model, unifies three advances—Interpretative Agency, Triadic STS Alignment, and a Recursive Learning Loop—that collectively extend TPB, Socio-Technical Systems Theory, and the Dynamic Capabilities View into the domain of algorithmic governance.

The STR model’s three components translate directly into sector-specific operational mandates, grounded in the empirical findings of Tables 12 and 13. First, in banking compliance systems, the Stability component requires that financial institutions implement model drift monitoring protocols calibrated to SHAP consistency metrics—ensuring that the fidelity of fraud-flag explanations remains stable across periods of market volatility and concept drift. A practical implementation vehicle is the quarterly ‘Interpretability Audit Report,’ benchmarked against internal Suspicious Activity Report (SAR) accuracy rates, which provides a structured, auditable record of explanation consistency over time [24,54,55].

Second, in audit processes, the Transparency component operationalises the EU AI Act’s ‘right to explanation’ in a concrete, audit-ready format: internal and external auditors should require that every AI-generated fraud flag be accompanied by a SHAP-based ‘Model Card’ as a precondition of SAR admissibility. This converts an abstract regulatory obligation into a documentable artefact that can be evaluated by non-technical compliance officers and legal review panels [7,57].

Third, in FinTech governance practices, the Reliability component mandates the institutionalisation of Human-in-the-Loop (HITL) oversight as a standard operating procedure within real-time fraud triage workflows. This should be supported by Cognitive Load Management interfaces—dashboard designs that reduce investigator ‘explanation fatigue’ by surfacing the three highest-weight SHAP features per alert rather than the full feature attribution vector [35,56]. Organisations that achieve this tripartite alignment—Stability in model outputs, Transparency in audit documentation, and Reliability in human oversight—are positioned not only to satisfy regulatory obligations but to operationalise explainability as a competitive differentiator: a ‘Transparent AI’ brand signal in an environment where institutional trust is a scarce and monetisable asset [27,53,184].

The generalisability of these findings is qualified by two scope limitations. First, although corpus-level quality was assured through the ABDC Journal Quality List filter, individual study-level risk of bias was not formally assessed using a quantitative instrument such as ROBIS; readers should interpret the ADO categorisations with this in mind. Second, as a systematic literature review limited by its 2015–2026 corpus window, the findings might not fully capture the latest deployments of real-time XAI in fast-evolving FinTech environments—a limitation that the Recursive Frontier research agenda in Table 15 is specifically designed to address. The methodological justification for the corpus design decisions (Scopus-only database, open-access filter, exclusion of IEEE/ACM) is reported in Section 2.3. Future research should extend this corpus cross-database and longitudinally to validate the STR Model’s recursive dynamics in live financial environments.

The study contributes to the FinTech literature by bridging the gap between technical explainability and regulatory governance, offering a structured framework that supports both academic research and practical implementation. Ultimately, explainable AI represents

a critical bridge between technological innovation and regulatory accountability in the evolving FinTech landscape.

Ultimately, this study positions explainable artificial intelligence as a critical enabler of trustworthy and scalable FinTech systems. By integrating behavioural theory with governance mechanisms, the proposed framework provides a foundation for future research and practical implementation in digitally transforming financial ecosystems.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/fintech5030060/s1>.

Author Contributions: Conceptualisation, D.G., P.C., K.S. and S.G.; methodology, D.G. and P.C.; formal analysis, D.G., P.C. and K.S.; investigation, D.G., P.C. and S.G.; data curation, D.G.; writing—original draft preparation, D.G. and P.C.; writing—review and editing, D.G., P.C., K.S. and S.G.; visualisation, D.G.; supervision, K.S. and S.G.; project administration, K.S. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable.

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

Abbreviations

The following abbreviations are used in this manuscript:

XAI	Explainable Artificial Intelligence
AI	Artificial Intelligence
FFD	Financial Fraud Detection
TPB	Theory of Planned Behaviour
PBC	Perceived Behavioural Control
ADO	Antecedents–Decisions–Outcomes
TCCM	Theory, Context, Characteristics, Methodology
STR	Stability–Transparency–Reliability
STS	Socio-Technical Systems
RBV	Resource-Based View
DCV	Dynamic Capabilities View
SLR	Systematic Literature Review
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SPAR-4-SLR	Structured Protocol for Reporting on Systematic Literature Reviews
ABDC	Australian Business Deans Council
GDPR	General Data Protection Regulation
AML	Anti-Money Laundering
SAR	Suspicious Activity Report
ESG	Environmental, Social, and Governance
HITL	Human-in-the-Loop
SHAP	SHapley Additive exPlanations
LIME	Local Interpretable Model-agnostic Explanations
GNN	Graph Neural Network
DNN	Deep Neural Network
RQ	Research Question
CRO	Chief Risk Officer
FinTech	Financial Technology

References

- Buchanan, B.G.; Wright, D. The Impact of Machine Learning on UK Financial Services. *Oxf. Rev. Econ. Policy* **2021**, *37*, 537–563. [CrossRef] [PubMed]
- Galaz, V.; Centeno, M.A.; Callahan, P.W.; Causevic, A.; Patterson, T.; Brass, I.; Baum, S.; Farber, D.; Fischer, J.; Garcia, D.; et al. Artificial Intelligence, Systemic Risks, and Sustainability. *Technol. Soc.* **2021**, *67*, 101741. [CrossRef]
- Ozturk, O. The Impact of AI on International Trade: Opportunities and Challenges. *Economies* **2024**, *12*, 298. [CrossRef]
- Bui, L.V. Legal and Regulatory Challenges in Addressing High-Tech Crimes in the Banking Sector: Developing a Cybersecurity Risk Assessment Framework for Vietnam. *J. Bank. Regul.* **2026**, *27*, 4. [CrossRef]
- Lokanan, M.E. Data Mining for Statistical Analysis of Money Laundering Transactions. *J. Money Laund. Control* **2019**, *22*, 753–763. [CrossRef]
- Mogre, R.; Talluri, S.S.; DAmico, F. A Decision Framework to Mitigate Supply Chain Risks: An Application in the Offshore-Wind Industry. *IEEE Trans. Eng. Manag.* **2016**, *63*, 316–325. [CrossRef]
- Turksen, U.; Benson, V.; Adamyk, B. Legal Implications of Automated Suspicious Transaction Monitoring: Enhancing Integrity of AI. *J. Bank. Regul.* **2024**, *25*, 359–377. [CrossRef]
- Ali, H.; Zafar, M.B.; Aysan, A.F. Generative AI in Finance: Replicability, Methodological Contingencies, and Future Research Directions. *Financ. Res. Lett.* **2025**, *86*, 108797. [CrossRef]
- Cordeiro, C.M.; Adomaitis, L.; Huang, L. The AI-Policy-Governance Nexus: How Regulation and AI Shift Corporate Governance toward Stakeholders. *Technol. Soc.* **2026**, *84*, 103117. [CrossRef]
- Mokoena, P.B. Harnessing Artificial Intelligence by Embedding Advanced Analytics and Modelling Techniques into Risk Management Processes. *Risk Manag. Insur. Rev.* **2025**, *28*, 207–231. [CrossRef]
- Farbmacher, H.; Löw, L.; Spindler, M. An Explainable Attention Network for Fraud Detection in Claims Management. *J. Econom.* **2022**, *228*, 244–258. [CrossRef]
- Shpachuk, V.; Markova, O.; Adamyk, B. AI-Driven Financial Fraud: Key Risks and Legal Protections for Financial Institutions. *J. Bank. Regul.* **2026**, *27*, 6. [CrossRef]
- Gomber, P.; Koch, J.-A.; Siering, M. Digital Finance and FinTech: Current Research and Future Research Directions. *J. Bus. Econ.* **2017**, *87*, 537–580. [CrossRef]
- United Nations Department of Economic and Social Affairs. *World Social Report 2023: Leaving No One Behind in an Ageing World*; World Social Report; United Nations: New York, NY, USA, 2023.
- IBM. What Is Fintech? Available online: <https://www.ibm.com/think/topics/fintech> (accessed on 25 May 2026).
- Arner, D.W. FinTech, RegTech, and the Reconceptualization of Financial Regulation. *Int. Law* **2017**, *37*, 371.
- Rodgers, W.; Hudson, R.; Economou, F. Modelling Credit and Investment Decisions Based on AI Algorithmic Behavioral Pathways. *Technol. Forecast. Soc. Change* **2023**, *191*, 122471. [CrossRef]
- Arnaboldi, M.; De Bruijn, H.; Steccolini, I.; Van Der Voort, H. On Humans, Algorithms and Data. *Qual. Res. Account. Manag.* **2022**, *19*, 241–254. [CrossRef]
- Long, P.; Solaimani, S. Beyond the Black Box: Operationalising Explicability in Artificial Intelligence for Financial Institutions. *Int. J. Bus. Inf. Syst.* **2025**, *49*, 10071822. [CrossRef]
- Hilb, M. Toward Artificial Governance? The Role of Artificial Intelligence in Shaping the Future of Corporate Governance. *J. Manag. Gov.* **2020**, *24*, 851–870. [CrossRef]
- Sætra, H.S. A Shallow Defence of a Technocracy of Artificial Intelligence: Examining the Political Harms of Algorithmic Governance in the Domain of Government. *Technol. Soc.* **2020**, *62*, 101283. [CrossRef] [PubMed]
- Cui, B.; Ge, L.; Grecov, P. Bond Defaults in China: Using Machine Learning to Make Predictions. *Int. Rev. Financ.* **2025**, *25*, e70010. [CrossRef]
- Rao, S.X.; Han, Z.; Yin, H.; Jiang, J.; Zhang, Z.; Zhao, Y.; Shan, Y. Fraud Detection at eBay. *Emerg. Mark. Rev.* **2025**, *66*, 101277. [CrossRef]
- Georgiou, K.; Yannacopoulos, A.N. Deep Neural Networks for Probability of Default Modelling. *J. Ind. Manag. Optim.* **2024**, *20*, 3647–3677. [CrossRef]
- Weber, P.; Carl, K.V.; Hinz, O. Applications of Explainable Artificial Intelligence in Finance—A Systematic Review of Finance, Information Systems, and Computer Science Literature. *Manag. Rev. Q.* **2024**, *74*, 867–907. [CrossRef]
- Schneider, M.; Brühl, R. Disentangling the Black Box around CEO and Financial Information-Based Accounting Fraud Detection: Machine Learning-Based Evidence from Publicly Listed U.S. Firms. *J. Bus. Econ.* **2023**, *93*, 1591–1628. [CrossRef]
- Akter, S.; Dwivedi, Y.K.; Biswas, K.; Michael, K.; Bandara, R.J.; Sajib, S. Addressing Algorithmic Bias in AI-Driven Customer Management. *J. Glob. Inf. Manag.* **2021**, *29*, 1–27. [CrossRef]
- Visani, G.; Bagli, E.; Chesani, F.; Poluzzi, A.; Capuzzo, D. Statistical Stability Indices for LIME: Obtaining Reliable Explanations for Machine Learning Models. *J. Oper. Res. Soc.* **2022**, *73*, 91–101. [CrossRef]

29. Hickman, E.; Petrin, M. Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective. *Eur. Bus. Organ. Law Rev.* **2021**, *22*, 593–625. [\[CrossRef\]](#)
30. Guo, T.; Liu, P.; Wang, C.; Xie, J.; Du, J.; Lim, M.K. Toward Sustainable Port-Hinterland Transportation: A Holistic Approach to Design Modal Shift Policy Mixes. *Transp. Res. Part Policy Pract.* **2023**, *174*, 103746. [\[CrossRef\]](#)
31. Rana, N.P.; Chatterjee, S.; Dwivedi, Y.K.; Akter, S. Understanding Dark Side of Artificial Intelligence (AI) Integrated Business Analytics: Assessing Firm's Operational Inefficiency and Competitiveness. *Eur. J. Inf. Syst.* **2022**, *31*, 364–387. [\[CrossRef\]](#)
32. Abedin, B. Managing the Tension between Opposing Effects of Explainability of Artificial Intelligence: A Contingency Theory Perspective. *Internet Res.* **2022**, *32*, 425–453. [\[CrossRef\]](#)
33. Chatti, H.; Argoubi, M. Artificial Intelligence in Knowledge Management: Identifying Intellectual Milestones and Emerging Domains. *Electron. J. Knowl. Manag.* **2025**, *23*, 122–148. [\[CrossRef\]](#)
34. Detthamrong, U.; Chansanam, W.; Boongoen, T.; Iam-On, N. Enhancing Fraud Detection in Banking Using Advanced Machine Learning Techniques. *Int. J. Econ. Financ. Issues* **2024**, *14*, 177–184. [\[CrossRef\]](#)
35. Ariza-Garzón, M.-J.; Arroyo, J.; Segovia-Vargas, M.-J.; Caparrini, A. Profit-Sensitive Machine Learning Classification with Explanations in Credit Risk: The Case of Small Businesses in Peer-to-Peer Lending. *Electron. Commer. Res. Appl.* **2024**, *67*, 101428. [\[CrossRef\]](#)
36. Zhao, Y. The Data Analysis of Enterprise Operational Risk Prediction Under Machine Learning: Innovations and Improvements in Corporate Law Risk Management Strategies. *J. Organ. End User Comput.* **2024**, *36*, 1–24. [\[CrossRef\]](#)
37. Kim, S.; Kim, C. Addressing Perceived Resistance to Biometric Security Systems in Airports: Exploration With a General Bayesian Network-Based Decision Support System. *J. Organ. End User Comput.* **2026**, *38*, 1–31. [\[CrossRef\]](#)
38. Brauner, P.; Glawe, F.; Liehner, G.L.; Vervier, L.; Ziefle, M. Mapping Public Perception of Artificial Intelligence: Expectations, Risk–Benefit Tradeoffs, and Value as Determinants for Societal Acceptance. *Technol. Forecast. Soc. Change* **2025**, *220*, 124304. [\[CrossRef\]](#)
39. Grimmelikhuijsen, S. Explaining Why the Computer Says No: Algorithmic Transparency Affects the Perceived Trustworthiness of Automated Decision-Making. *Public Adm. Rev.* **2023**, *83*, 241–262. [\[CrossRef\]](#)
40. Riedl, R. Is Trust in Artificial Intelligence Systems Related to User Personality? Review of Empirical Evidence and Future Research Directions. *Electron. Mark.* **2022**, *32*, 2021–2051. [\[CrossRef\]](#)
41. Kim, T.W.; Routledge, B.R. Why a Right to an Explanation of Algorithmic Decision-Making Should Exist: A Trust-Based Approach. *Bus. Ethics Q.* **2022**, *32*, 75–102. [\[CrossRef\]](#)
42. Gupta, A.; Dwivedi, D.N.; Shah, J.; Jain, A. Data Quality Issues Leading to Sub Optimal Machine Learning for Money Laundering Models. *J. Money Laund. Control* **2022**, *25*, 551–555. [\[CrossRef\]](#)
43. Baesens, B.; Höppner, S.; Verdonck, T. Data Engineering for Fraud Detection. *Decis. Support Syst.* **2021**, *150*, 113492. [\[CrossRef\]](#)
44. Gerlach, J.; Hoppe, P.; Jagels, S.; Licker, L.; Breitner, M.H. Decision Support for Efficient XAI Services—A Morphological Analysis, Business Model Archetypes, and a Decision Tree. *Electron. Mark.* **2022**, *32*, 2139–2158. [\[CrossRef\]](#)
45. Shen, F.; Zhao, X.; Kou, G. Three-Stage Reject Inference Learning Framework for Credit Scoring Using Unsupervised Transfer Learning and Three-Way Decision Theory. *Decis. Support Syst.* **2020**, *137*, 113366. [\[CrossRef\]](#)
46. Nadeem, W.; Ashraf, A.R.; Khan, H.; Kumar, V. Impact of AI Strategies on Climate-Change Performance: Responsible AI and Crisis Management Perspectives. *Technovation* **2026**, *150*, 103390. [\[CrossRef\]](#)
47. Rodríguez-Espíndola, O.; Chowdhury, S.; Dey, P.K.; Albores, P.; Emrouznejad, A. Analysis of the Adoption of Emergent Technologies for Risk Management in the Era of Digital Manufacturing. *Technol. Forecast. Soc. Change* **2022**, *178*, 121562. [\[CrossRef\]](#)
48. Berman, A.; De Fine Licht, K.; Carlsson, V. Trustworthy AI in the Public Sector: An Empirical Analysis of a Swedish Labor Market Decision-Support System. *Technol. Soc.* **2024**, *76*, 102471. [\[CrossRef\]](#)
49. Aldemir, C.; Uçma Uysal, T. Artificial Intelligence for Financial Accountability and Governance in the Public Sector: Strategic Opportunities and Challenges. *Adm. Sci.* **2025**, *15*, 58. [\[CrossRef\]](#)
50. Coita, I.-F.; Belbe, S.S.; Mare, C.C.; Osterrieder, J.; Hopp, C. Modelling Taxpayers' Behaviour Based on Prediction of Trust Using Sentiment Analysis. *Financ. Res. Lett.* **2023**, *58*, 104549. [\[CrossRef\]](#)
51. Tenzer, M. Social Landscape Characterisation: A People-Centred, Place-Based Approach to Inclusive and Transparent Heritage and Landscape Management. *Int. J. Herit. Stud.* **2024**, *30*, 269–284. [\[CrossRef\]](#)
52. Hor, T.S.C.; Fong, L.; Wynne, K.; Verhoeven, B. Easing the Cognitive Load of General Practitioners: AI Design Principles for Future-Ready Healthcare. *Technovation* **2025**, *142*, 103208. [\[CrossRef\]](#)
53. Nyqvist, R.; Peltokorpi, A.; Seppänen, O. Can ChatGPT Exceed Humans in Construction Project Risk Management? *Eng. Constr. Archit. Manag.* **2024**, *31*, 223–243. [\[CrossRef\]](#)
54. Poszler, F.; Lange, B. The Impact of Intelligent Decision-Support Systems on Humans' Ethical Decision-Making: A Systematic Literature Review and an Integrated Framework. *Technol. Forecast. Soc. Change* **2024**, *204*, 123403. [\[CrossRef\]](#)
55. Glaser, V.L.; Sloan, J.; Gehman, J. Organizations as Algorithms: A New Metaphor for Advancing Management Theory. *J. Manag. Stud.* **2024**, *61*, 2748–2769. [\[CrossRef\]](#)

56. Hülter, S.M.; Ertel, C.; Heidemann, A. Exploring the Individual Adoption of Human Resource Analytics: Behavioural Beliefs and the Role of Machine Learning Characteristics. *Technol. Forecast. Soc. Change* **2024**, *208*, 123709. [\[CrossRef\]](#)
57. Pande, D.; Taeihagh, A. A Governance Perspective on User Acceptance of Autonomous Systems in Singapore. *Technol. Soc.* **2024**, *77*, 102580. [\[CrossRef\]](#)
58. Liu, J.; Zhang, X.; Xiong, H. Credit Risk Prediction Based on Causal Machine Learning: Bayesian Network Learning, Default Inference, and Interpretation. *J. Forecast.* **2024**, *43*, 1625–1660. [\[CrossRef\]](#)
59. Liu, X.; Lu, J.; Cheng, Z.; Ma, X. A Dynamic Bayesian Network-Based Real-Time Crash Prediction Model for Urban Elevated Expressway. *J. Adv. Transp.* **2021**, *2021*, 5569143. [\[CrossRef\]](#)
60. Zacharias, J.; Von Zahn, M.; Chen, J.; Hinz, O. Designing a Feature Selection Method Based on Explainable Artificial Intelligence. *Electron. Mark.* **2022**, *32*, 2159–2184. [\[CrossRef\]](#)
61. Peeters, B. AI-Based Profiling by Tax Authorities: Exploring GDPR Constraints and Explainability. *EC Tax Rev.* **2025**, *34*, 182–185. [\[CrossRef\]](#)
62. Tranfield, D.; Denyer, D.; Smart, P. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *Br. J. Manag.* **2003**, *14*, 207–222. [\[CrossRef\]](#)
63. Snyder, H. Literature Review as a Research Methodology: An Overview and Guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [\[CrossRef\]](#)
64. Webster, J.; Watson, R. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Q.* **2002**, *26*, xiii–xxiii. [\[CrossRef\]](#)
65. Page, M.; McKenzie, J.; Bossuyt, P.; Boutron, I.; Hoffmann, T.; Mulrow, C.; Shamseer, L.; Tetzlaff, J.; Akl, E.; Brennan, S.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Paul, J.; Lim, W.M.; O’Cass, A.; Hao, A.W.; Bresciani, S. Scientific Procedures and Rationales for Systematic Literature Reviews (SPAR-4-SLR). *Int. J. Consum. Stud.* **2021**, *45*, O1–O16. [\[CrossRef\]](#)
67. Okoli, C.; Schabram, K. A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electron. J.* **2010**, *10*. [\[CrossRef\]](#)
68. Kitchenham, B.; Charters, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; Keele University: Keele, UK; University of Durham: Durham, UK, 2007.
69. Denyer, D.; Tranfield, D. Producing a Systematic Review. In *The SAGE Handbook of Organizational Research Methods*; SAGE Publications Ltd.: London, UK, 2009; pp. 671–689.
70. Barredo Arrieta, A.; Díaz-Rodríguez, N.; Del Ser, J.; Benetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [\[CrossRef\]](#)
71. Ali, S.; Abuhmed, T.; El-Sappagh, S.; Muhammad, K.; Alonso-Moral, J.M.; Confalonieri, R.; Guidotti, R.; Del Ser, J.; Díaz-Rodríguez, N.; Herrera, F. Explainable Artificial Intelligence (XAI): What We Know and What Is Left to Attain Trustworthy Artificial Intelligence. *Inf. Fusion* **2023**, *99*, 101805. [\[CrossRef\]](#)
72. Bennani, M.H.; Driss, H. Rethinking Systematic Literature Reviews: A Critical and Evolving Perspective on the SPAR-4-SLR Protocol. *Rev. Int. Des Sci. De Gest.* **2025**, *8*, 519–539.
73. Sauer, P.C.; Seuring, S. How to Conduct Systematic Literature Reviews in Management Research: A Guide in 6 Steps and 14 Decisions. *Rev. Manag. Sci.* **2023**, *17*, 1899–1933. [\[CrossRef\]](#)
74. Olan, F.; Spanaki, K.; Ahmed, W.; Zhao, G. Enabling Explainable Artificial Intelligence Capabilities in Supply Chain Decision Support Making. *Prod. Plan. Control* **2025**, *36*, 808–819. [\[CrossRef\]](#)
75. Hidalgo, P.; Rodriguez, D. An Explainable Multi-Task Similarity Measure: Integrating Accumulated Local Effects and Weighted Fréchet Distance. *Knowl.-Based Syst.* **2025**, *329*, 114384. [\[CrossRef\]](#)
76. Mongeon, P.; Paul-Hus, A. The Journal Coverage of Web of Science and Scopus: A Comparative Analysis. *Scientometrics* **2016**, *106*, 213–228. [\[CrossRef\]](#)
77. Wilson, C.; Van Der Velden, M. Sustainable AI: An Integrated Model to Guide Public Sector Decision-Making. *Technol. Soc.* **2022**, *68*, 101926. [\[CrossRef\]](#)
78. Feldkamp, T.; Langer, M.; Wies, L.; König, C.J. Justice, Trust, and Moral Judgements When Personnel Selection Is Supported by Algorithms. *Eur. J. Work Organ. Psychol.* **2024**, *33*, 130–145. [\[CrossRef\]](#)
79. Liu, Z.; Liang, H. Are Credit Scores Gender-Neutral? Evidence of Mis-Calibration from Alternative and Traditional Borrowing Data. *J. Behav. Exp. Financ.* **2025**, *47*, 101081. [\[CrossRef\]](#)
80. Moreira, L.; Santos, I.L.D.; Gonzalez, P.H. Portfolio Optimization for Pension Purposes: Literature Review. *J. Econ. Surv.* **2026**, *40*, 45–72. [\[CrossRef\]](#)
81. Yang, S.; Wu, H. The Global Organizational Behavior Analysis for Financial Risk Management Utilizing Artificial Intelligence. *J. Glob. Inf. Manag.* **2021**, *30*, 1–24. [\[CrossRef\]](#)

82. Ghobakhloo, M.; Fathi, M.; Iranmanesh, M.; Vilkas, M.; Grybauskas, A.; Amran, A. Generative Artificial Intelligence in Manufacturing: Opportunities for Actualizing Industry 5.0 Sustainability Goals. *J. Manuf. Technol. Manag.* **2024**, *35*, 94–121. [\[CrossRef\]](#)
83. Zhu, J. Application of Artificial Intelligence Data Mining Algorithm in Enterprise Management Risk Assessment. *Int. J. Inf. Syst. Supply Chain Manag.* **2024**, *17*, 1–19. [\[CrossRef\]](#)
84. Martinez-Torres, M.R.; Toral, S.L. A Machine Learning Approach for the Identification of the Deceptive Reviews in the Hospitality Sector Using Unique Attributes and Sentiment Orientation. *Tour. Manag.* **2019**, *75*, 393–403. [\[CrossRef\]](#)
85. Wruck, S.; Vis, I.F.A.; Boter, J. Risk Control for Staff Planning in E-Commerce Warehouses. *Int. J. Prod. Res.* **2017**, *55*, 6453–6469. [\[CrossRef\]](#)
86. De Blasio, G.; D'Ignazio, A.; Letta, M. Gotham City. Predicting 'Corrupted' Municipalities with Machine Learning. *Technol. Forecast. Soc. Change* **2022**, *184*, 122016. [\[CrossRef\]](#)
87. Aldboush, H.H.H.; Ferdous, M. Building Trust in Fintech: An Analysis of Ethical and Privacy Considerations in the Intersection of Big Data, AI, and Customer Trust. *Int. J. Financ. Stud.* **2023**, *11*, 90. [\[CrossRef\]](#)
88. Singh, C.; Lin, W. Can Artificial Intelligence, RegTech and CharityTech Provide Effective Solutions for Anti-Money Laundering and Counter-Terror Financing Initiatives in Charitable Fundraising. *J. Money Laund. Control* **2021**, *24*, 464–482. [\[CrossRef\]](#)
89. Hajiali, M.; Teimoury, E.; Rabiee, M.; Delen, D. An Interactive Decision Support System for Real-Time Ambulance Relocation with Priority Guidelines. *Decis. Support Syst.* **2022**, *155*, 113712. [\[CrossRef\]](#)
90. Selten, F.; Robeer, M.; Grimmelikhuijsen, S. 'Just like I Thought': Street-level Bureaucrats Trust AI Recommendations If They Confirm Their Professional Judgment. *Public Adm. Rev.* **2023**, *83*, 263–278. [\[CrossRef\]](#)
91. Jeremiah, F. The Human-AI Dyad: Navigating the New Frontier of Entrepreneurial Discourse. *Futures* **2025**, *166*, 103529. [\[CrossRef\]](#)
92. Zeng, S.; Dai, S. Synergizing Domain Knowledge and Machine Learning: Intelligent Early Fraud Detection Enhanced by Earnings Management Analysis. *Int. Rev. Financ.* **2025**, *25*, e70021. [\[CrossRef\]](#)
93. Chindasombatcharoen, P.; Chatjuthamard, P.; Jiraporn, P.; Wongsinhirun, N. Climate Change Exposure and the Takeover Market: A Text-Based Analysis. *Bus. Strategy Environ.* **2024**, *33*, 8587–8592. [\[CrossRef\]](#)
94. Eroğlu, M.; Karatepe Kaya, M. Impact of Artificial Intelligence on Corporate Board Diversity Policies and Regulations. *Eur. Bus. Organ. Law Rev.* **2022**, *23*, 541–572. [\[CrossRef\]](#)
95. Chamochumbi Diaz, G.D.; Palazzi, F.; Sorini, L. Creditworthiness of Small and Medium Enterprises: A Fuzzy Decision-Making Approach. *Comput. Manag. Sci.* **2026**, *23*, 2. [\[CrossRef\]](#)
96. Pocher, N.; Zichichi, M.; Merizzi, F.; Shafiq, M.Z.; Ferretti, S. Detecting Anomalous Cryptocurrency Transactions: An AML/CFT Application of Machine Learning-Based Forensics. *Electron. Mark.* **2023**, *33*, 37. [\[CrossRef\]](#)
97. Matthews, J.; Love, P.E.D.; Porter, S.; Fang, W. Curating a Domain Ontology for Rework in Construction: Challenges and Learnings from Practice. *Prod. Plan. Control* **2024**, *35*, 2068–2083. [\[CrossRef\]](#)
98. Rožanec, J.M.; Novalija, I.; Zajec, P.; Kenda, K.; Tavakoli Ghinani, H.; Suh, S.; Bian, S.; Veliou, E.; Papamartzivanos, D.; Giannetsos, T.; et al. Human-Centric Artificial Intelligence Architecture for Industry 5.0 Applications. *Int. J. Prod. Res.* **2023**, *61*, 6847–6872. [\[CrossRef\]](#)
99. Osei-Assibey Bonsu, M.; Wang, Y.; Guo, Y. Does Fintech Lead to Better Accounting Practices? Empirical Evidence. *Account. Res. J.* **2023**, *36*, 129–147. [\[CrossRef\]](#)
100. Pahari, S.; Polisetty, A.; Sharma, S.; Jha, R.; Chakraborty, D. Adoption of AI in the Banking Industry: A Case Study on Indian Banks. *Indian J. Mark.* **2023**, *53*, 26. [\[CrossRef\]](#)
101. Kim, D.H.; Vanheusden, F.J.; Kim, A. Forecasting Cryptocurrency Markets Using Recurrence and Time-Frequency Analysis-Based Machine Learning Algorithms. *Financ. Res. Lett.* **2025**, *85*, 108268. [\[CrossRef\]](#)
102. Yu, X.; Chen, Y.; Zhao, Q.; Li, H. Research and Analysis of Trust and Control in Human-AI Interaction for Decision-Making Systems in Optimizing Public Health Service. *J. Organ. End User Comput.* **2025**, *37*, 1–15. [\[CrossRef\]](#)
103. Hasan, M.M.; Taylor, G.; Richardson, G. Brand Capital and Stock Price Crash Risk. *Manag. Sci.* **2022**, *68*, 7221–7247. [\[CrossRef\]](#)
104. Hyde, S.J.; Bachura, E.; Bundy, J.; Gretz, R.T.; Sanders, W.G. The Tangled Webs We Weave: Examining the Effects of CEO Deception on Analyst Recommendations. *Strateg. Manag. J.* **2024**, *45*, 66–112. [\[CrossRef\]](#)
105. Düsterhöft, M.; Schiemann, F.; Walther, T. Let's Talk about Risk! Stock Market Effects of Risk Disclosure for European Energy Utilities. *Energy Econ.* **2023**, *125*, 106794. [\[CrossRef\]](#)
106. Li, W.; Liu, X.; Su, J.; Cui, T. Advancing Financial Risk Management: A Transparent Framework for Effective Fraud Detection. *Financ. Res. Lett.* **2025**, *75*, 106865. [\[CrossRef\]](#)
107. Chen, S. Credit Scoring Prediction for Small and Medium-Sized Enterprises Based on Alternative Data and Gradient Boosting Algorithms. *Inf. Resour. Manag. J.* **2026**, *39*, 1–19. [\[CrossRef\]](#)
108. Wen, W.; Han, X. An Introduction of Transaction Session-induced Security Scheme Using Blockchain Technology: Understanding the Features of Internet of Things-Based Financial Security Systems. *Manag. Decis. Econ.* **2024**, *45*, 1817–1834. [\[CrossRef\]](#)

109. Chen, M.; Wei, Y.; Wang, S. Past and Future of Cryptocurrencies: A Survey Using Bibliometric Methods. *J. Econ. Surv.* **2026**, *40*, 20–44. [\[CrossRef\]](#)
110. Rafi-Ul-Shan, P.M.; Bashiri, M.; Kamal, M.M.; Mangla, S.K.; Tjahjono, B. An Analysis of Fuzzy Group Decision Making to Adopt Emerging Technologies for Fashion Supply Chain Risk Management. *IEEE Trans. Eng. Manag.* **2024**, *71*, 8469–8487. [\[CrossRef\]](#)
111. Castañé, G.; Dolgui, A.; Kousi, N.; Meyers, B.; Thevenin, S.; Vyhmeister, E.; Östberg, P.-O. The ASSISTANT Project: AI for High Level Decisions in Manufacturing. *Int. J. Prod. Res.* **2023**, *61*, 2288–2306. [\[CrossRef\]](#)
112. Liu, C.; Ryan, D.; Lin, G.; Xu, C. No Rose without a Thorn: Corporate Teamwork Culture and Financial Statement Misconduct. *J. Behav. Exp. Financ.* **2023**, *37*, 100786. [\[CrossRef\]](#)
113. Pessach, D.; Singer, G.; Avrahami, D.; Chalutz Ben-Gal, H.; Shmueli, E.; Ben-Gal, I. Employees Recruitment: A Prescriptive Analytics Approach via Machine Learning and Mathematical Programming. *Decis. Support Syst.* **2020**, *134*, 113290. [\[CrossRef\]](#) [\[PubMed\]](#)
114. Salas-Molina, F. Selecting the Best Risk Measure in Multiobjective Cash Management. *Int. Trans. Oper. Res.* **2019**, *26*, 929–945. [\[CrossRef\]](#)
115. Oyenubi, A.; Kollamparambil, U. Does Noncompliance with COVID-19 Regulations Impact the Depressive Symptoms of Others? *Econ. Model.* **2023**, *120*, 106191. [\[CrossRef\]](#) [\[PubMed\]](#)
116. Nguyen Thanh, C.; Phan Huy, T. Predicting Financial Reports Fraud by Machine Learning: The Proxy of Auditor Opinions. *Cogent Bus. Manag.* **2025**, *12*, 2510556. [\[CrossRef\]](#)
117. Wang, L.; Yu, Z.; Ma, J.; Chen, X.; Wu, C. A Two-Stage Interpretable Model to Explain Classifier in Credit Risk Prediction. *J. Forecast.* **2025**, *44*, 2132–2150. [\[CrossRef\]](#)
118. Banulescu-Radu, D.; Yankol-Schalck, M. Practical Guideline to Efficiently Detect Insurance Fraud in the Era of Machine Learning: A Household Insurance Case. *J. Risk Insur.* **2024**, *91*, 867–913. [\[CrossRef\]](#)
119. Vanini, P.; Rossi, S.; Zvizdic, E.; Domenig, T. Online Payment Fraud: From Anomaly Detection to Risk Management. *Financ. Innov.* **2023**, *9*, 66. [\[CrossRef\]](#)
120. Feng, T.; Xu, J.; Zhou, Z.; Luo, Y. How Green Credit Policy Affects Commercial Banks' Credit Risk?: Evidence and Federated Learning-Based Modeling From 26 Listed Commercial Banks in China. *J. Cases Inf. Technol.* **2023**, *26*, 1–21. [\[CrossRef\]](#)
121. Niu, S.; Yin, Q.; Ma, J.; Song, Y.; Xu, Y.; Bai, L.; Pan, W.; Yang, X. Enhancing Healthcare Decision Support through Explainable AI Models for Risk Prediction. *Decis. Support Syst.* **2024**, *181*, 114228. [\[CrossRef\]](#)
122. Vandervorst, F.; Verbeke, W.; Verdonck, T. Data Misrepresentation Detection for Insurance Underwriting Fraud Prevention. *Decis. Support Syst.* **2022**, *159*, 113798. [\[CrossRef\]](#)
123. Lu, Z.; Li, H.; Wu, J. Exploring the Impact of Financial Literacy on Predicting Credit Default among Farmers: An Analysis Using a Hybrid Machine Learning Model. *Borsa Istanbul. Rev.* **2024**, *24*, 352–362. [\[CrossRef\]](#)
124. Biswas, B.; Mukhopadhyay, A.; Kumar, A.; Delen, D. A Hybrid Framework Using Explainable AI (XAI) in Cyber-Risk Management for Defence and Recovery against Phishing Attacks. *Decis. Support Syst.* **2024**, *177*, 114102. [\[CrossRef\]](#)
125. Kong, L.; Brintrup, A. A Hierarchical Bayesian Model for Payment Delay Prediction in Supply Chain Financing. *Int. J. Prod. Res.* **2026**, *64*, 168–191. [\[CrossRef\]](#)
126. Luna, M.; Pérez-Mon, O.; Becker, J.L. Forecasting and Managing Price Volatility in Salmon Production: A Hybrid System Using Conformal Prediction and Dynamic Hedging. *Int. J. Prod. Econ.* **2026**, *291*, 109726. [\[CrossRef\]](#)
127. Sadeh, H.; Mirarchi, C.; Shahbodaghlou, F.; Pavan, A. Predicting the Trends and Cost Impact of COVID-19 OSHA Citations on US Construction Contractors Using Machine Learning and Simulation. *Eng. Constr. Archit. Manag.* **2023**, *30*, 3461–3479. [\[CrossRef\]](#)
128. Feng, Y.; Wang, X.; Chen, Q.; Yang, Z.; Wang, J.; Li, H.; Xia, G.; Liu, Z. Prediction of the Severity of Marine Accidents Using Improved Machine Learning. *Transp. Res. Part E Logist. Transp. Rev.* **2024**, *188*, 103647. [\[CrossRef\]](#)
129. Hosseini Shekarabi, S.A.; Kiani Mavi, R.; Kiani Mavi, N.; Macau, F.R.; Arisian, S. (Sean) A Novel Robust Optimization Approach for Supply Chain Resilience: The Role of Flexibility and Collaboration. *Int. J. Prod. Econ.* **2025**, *287*, 109686. [\[CrossRef\]](#)
130. Dai, S.; Zhang, J.; Huang, Z.; Zeng, S. Fire Prediction and Risk Identification With Interpretable Machine Learning. *J. Forecast.* **2025**, *44*, 1699–1715. [\[CrossRef\]](#)
131. Martin, H.; James, J.; Chadee, A. Exploring Large Language Model AI Tools in Construction Project Risk Assessment: Chat GPT Limitations in Risk Identification, Mitigation Strategies, and User Experience. *J. Constr. Eng. Manag.* **2025**, *151*, 04025119. [\[CrossRef\]](#)
132. Popa, C.; Stefanov, O.; Goia, I. Multimodal Livestock Operations Analysis Using Business Process Modeling: A Case Study of Romanian Black Sea Ports. *Economies* **2025**, *13*, 69. [\[CrossRef\]](#)
133. Ghamarimajd, Z.; Ghanbaripour, A.; Tumpa, R.J.; Watanabe, T.; Mbachu, J.; Skitmore, M. Application of Systems Thinking and System Dynamics in Managing Risks and Stakeholders in Construction Projects: A Systematic Literature Review. *Syst. Res. Behav. Sci.* **2025**, *42*, 1465–1479. [\[CrossRef\]](#)
134. Mohamed, O.Y.; Binti Jamaludin, N.F. Development of an Integrated ESG and Climate Risk Assessment Framework for Semiconductors Industries in Malaysia. *Sustain. Futur.* **2025**, *10*, 101052. [\[CrossRef\]](#)

135. Pietersen, L.-A.; Rudman, R. Data-Related Risks for the Use of Machine Learning in Retail Customer Demand Forecasting. *S. Afr. J. Bus. Manag.* **2025**, *56*, 13. [\[CrossRef\]](#)
136. Ariza-Garzón, M.-J.; Camacho-Miñano, M.-D.-M.; Segovia-Vargas, M.-J.; Arroyo, J. Risk-Return Modelling in the P2p Lending Market: Trends, Gaps, Recommendations and Future Directions. *Electron. Commer. Res. Appl.* **2021**, *49*, 101079. [\[CrossRef\]](#)
137. West, J.; Bhattacharya, M. Intelligent Financial Fraud Detection: A Comprehensive Review. *Comput. Secur.* **2016**, *57*, 47–66. [\[CrossRef\]](#)
138. Boiko, R.; Butenko, D.; Frolov, A.; Moisiakha, A.; Rudevskaya, V. Evaluating Modern Quantitative Methods for Investment Portfolio Management under Market Uncertainty. *J. Appl. Econ. Sci. JAES* **2025**, *20*, 427. [\[CrossRef\]](#) [\[PubMed\]](#)
139. Ben Jabeur, S.; Bakkar, Y.; Cepni, O. Do Global COVOL and Geopolitical Risks Affect Clean Energy Prices? Evidence from Explainable Artificial Intelligence Models. *Energy Econ.* **2025**, *141*, 108112. [\[CrossRef\]](#)
140. Zhang, C.; Wang, Y.; Zhang, J. Risk Assessment of Live-Streaming Marketing Based on Hesitant Fuzzy Multi-Attribute Group Decision-Making Method. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 120. [\[CrossRef\]](#)
141. Zhang, Q.; Ma, Z.; Wu, Y.; Liu, Y.; Qu, X. Quantifying Variable Contributions to Bus Operation Delays Considering Causal Relationships. *Transp. Res. Part E Logist. Transp. Rev.* **2025**, *194*, 103881. [\[CrossRef\]](#)
142. Manimuthu, A.; Venkatesh, V.G.; Shi, Y.; Sreedharan, V.R.; Koh, S.C.L. Design and Development of Automobile Assembly Model Using Federated Artificial Intelligence with Smart Contract. *Int. J. Prod. Res.* **2022**, *60*, 111–135. [\[CrossRef\]](#)
143. Lopez-Martinez-Carrasco, A.; Juarez, J.M.; Campos, M.; Canovas-Segura, B. A Methodology Based on Trace-Based Clustering for Patient Phenotyping. *Knowl.-Based Syst.* **2021**, *232*, 107469. [\[CrossRef\]](#)
144. Paul, J.; Benito, G. A Review of Research on Outward Foreign Direct Investment from Emerging Countries, Including China: What Do We Know, How Do We Know and Where Should We Be Heading? *Asia Pac. Bus. Rev.* **2017**, *24*, 90–115. [\[CrossRef\]](#)
145. Paul, J.; Rosado-Serrano, A. Gradual Internationalization vs Born-Global/International New Venture Models: A Review and Research Agenda. *Int. Mark. Rev.* **2019**, *36*, 830–858. [\[CrossRef\]](#)
146. Ambika, A.; Shin, H.; Jain, V. Immersive Technologies and Consumer Behavior: A Systematic Review of Two Decades of Research. *Aust. J. Manag.* **2025**, *50*, 55–79. [\[CrossRef\]](#)
147. Thomas, A.; Gupta, V. Tacit Knowledge in Organizations: Bibliometrics and a Framework-Based Systematic Review of Antecedents, Outcomes, Theories, Methods and Future Directions. *J. Knowl. Manag.* **2022**, *26*, 1014–1041. [\[CrossRef\]](#)
148. Ajzen, I. From Intentions to Actions: A Theory of Planned Behavior. In *Action Control: From Cognition to Behavior*; Kuhl, J., Beckmann, J., Eds.; Springer: Berlin/Heidelberg, Germany, 1985; pp. 11–39.
149. Ajzen, I. The Theory of Planned Behavior. *Theor. Cogn. Self-Regul.* **1991**, *50*, 179–211. [\[CrossRef\]](#)
150. Ajzen, I. The Theory of Planned Behaviour: Reactions and Reflections. *Psychol. Health* **2011**, *26*, 1113–1127. [\[CrossRef\]](#) [\[PubMed\]](#)
151. Dwivedi, Y.; Hughes, L.; Ismagilova, E.; Aarts, G.; Coombs, C.; Crick, T.; Duan, Y.; Dwivedi, R.; Edwards, J.; Eirug, A.; et al. Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy. *Int. J. Inf. Manag.* **2019**, *57*, 101994. [\[CrossRef\]](#)
152. Al-Emran, M.; Mezhyuev, V.; Kamaludin, A. Technology Acceptance Model in M-Learning Context: A Systematic Review. *Comput. Educ.* **2018**, *125*, 389–412. [\[CrossRef\]](#)
153. De Oliveira, J.A.P.; Wanke, P.; Antunes, J.; Tan, Y. Unveiling the Impact of Information Vagueness on Carbon Emission Inventories Using Fuzzy Sets. *Energy Econ.* **2025**, *148*, 108672. [\[CrossRef\]](#)
154. Gangwar, H.; Date, H.; Ramaswamy, R. Understanding Determinants of Cloud Computing Adoption Using an Integrated TAM-TOE Model. *J. Enterp. Inf. Manag.* **2015**, *28*, 107–130. [\[CrossRef\]](#)
155. Xie, Z.; He, T.; Tian, S.; Fu, Y.; Zhou, J.; Chen, D. Joint Gaussian Mixture Model for Versatile Deep Visual Model Explanation. *Knowl.-Based Syst.* **2023**, *280*, 110989. [\[CrossRef\]](#)
156. Aljunaid, S.K.; Almheiri, S.J.; Dawood, H.; Khan, M.A. Secure and Transparent Banking: Explainable AI-Driven Federated Learning Model for Financial Fraud Detection. *J. Risk Financ. Manag.* **2025**, *18*, 179. [\[CrossRef\]](#)
157. Ghosh, I.; Alfaro-Cortés, E.; Gámez, M.; García-Rubio, N. Prediction and Interpretation of Daily NFT and DeFi Prices Dynamics: Inspection through Ensemble Machine Learning & XAI. *Int. Rev. Financ. Anal.* **2023**, *87*, 102558. [\[CrossRef\]](#)
158. Oprea, S.-V.; Băra, A. Customer-Centric Decision-Making with XAI and Counterfactual Explanations for Churn Mitigation. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 129. [\[CrossRef\]](#)
159. Almaqtari, F.A. The Role of IT Governance in the Integration of AI in Accounting and Auditing Operations. *Economies* **2024**, *12*, 199. [\[CrossRef\]](#)
160. Zebec, A.; Indihar Štemberger, M. Creating AI Business Value through BPM Capabilities. *Bus. Process Manag. J.* **2024**, *30*, 1–26. [\[CrossRef\]](#)
161. Boob-Engel, K. Is AI as Trustworthy as Human Experts? The Role of Advice Source, Technology Knowledge, and Perceptions of AI's Impact in Organizational Decision-Making. *J. Decis. Syst.* **2025**, *34*, 2594620. [\[CrossRef\]](#)
162. Rodgers, W.; Murray, J.M.; Stefanidis, A.; Degbey, W.Y.; Tarba, S.Y. An Artificial Intelligence Algorithmic Approach to Ethical Decision-Making in Human Resource Management Processes. *Hum. Resour. Manag. Rev.* **2023**, *33*, 100925. [\[CrossRef\]](#)

163. Loecher, M. Debiasing SHAP Scores in Random Forests. *ASTA Adv. Stat. Anal.* **2024**, *108*, 427–440. [\[CrossRef\]](#)
164. Xie, S. Improving Explainability of Major Risk Factors in Artificial Neural Networks for Auto Insurance Rate Regulation. *Risks* **2021**, *9*, 126. [\[CrossRef\]](#)
165. Gosiewska, A.; Kozak, A.; Biecek, P. Simpler Is Better: Lifting Interpretability-Performance Trade-off via Automated Feature Engineering. *Decis. Support Syst.* **2021**, *150*, 113556. [\[CrossRef\]](#)
166. Lorenz, F.; Willwersch, J.; Cajias, M.; Fuerst, F. Interpretable Machine Learning for Real Estate Market Analysis. *Real Estate Econ.* **2023**, *51*, 1178–1208. [\[CrossRef\]](#)
167. Peukert, C.; Bechtold, S.; Batikas, M.; Kretschmer, T. Regulatory Spillovers and Data Governance: Evidence from the GDPR. *Mark. Sci.* **2022**, *41*, 746–768. [\[CrossRef\]](#)
168. Brotcke, L. Time to Assess Bias in Machine Learning Models for Credit Decisions. *J. Risk Financ. Manag.* **2022**, *15*, 165. [\[CrossRef\]](#)
169. Zimmermann, R.; Mora, D.; Cirqueira, D.; Helfert, M.; Bezbradica, M.; Werth, D.; Weitzl, W.J.; Riedl, R.; Auinger, A. Enhancing Brick-and-Mortar Store Shopping Experience with an Augmented Reality Shopping Assistant Application Using Personalized Recommendations and Explainable Artificial Intelligence. *J. Res. Interact. Mark.* **2023**, *17*, 273–298. [\[CrossRef\]](#)
170. Tiukhova, E.; Vemuri, P.; Flores, N.L.; Islind, A.S.; Óskarsdóttir, M.; Poelmans, S.; Baesens, B.; Snoeck, M. Explainable Learning Analytics: Assessing the Stability of Student Success Prediction Models by Means of Explainable AI. *Decis. Support Syst.* **2024**, *182*, 114229. [\[CrossRef\]](#)
171. Senoner, J.; Netland, T.; Feuerriegel, S. Using Explainable Artificial Intelligence to Improve Process Quality: Evidence from Semiconductor Manufacturing. *Manag. Sci.* **2022**, *68*, 5704–5723. [\[CrossRef\]](#)
172. Nimmy, S.F.; Hussain, O.K.; Chakraborty, R.K.; Hussain, F.K.; Saberi, M. An Optimized Belief-Rule-Based (BRB) Approach to Ensure the Trustworthiness of Interpreted Time-Series Decisions. *Knowl.-Based Syst.* **2023**, *271*, 110552. [\[CrossRef\]](#)
173. Zhong, C.; Goel, S. Transparent AI in Auditing through Explainable AI. *Curr. Issues Audit.* **2024**, *18*, A1–A14. [\[CrossRef\]](#)
174. Agbabiaka, O.; Ojo, A.; Connolly, N. Requirements for Trustworthy AI-Enabled Automated Decision-Making in the Public Sector: A Systematic Review. *Technol. Forecast. Soc. Change* **2025**, *215*, 124076. [\[CrossRef\]](#)
175. Zhu, L.; Chen, P.; Dong, D.; Wang, Z. Can Artificial Intelligence Enable the Government to Respond More Effectively to Major Public Health Emergencies?—Taking the Prevention and Control of Covid-19 in China as an Example. *Socio-Econ. Plan. Sci.* **2022**, *80*, 101029. [\[CrossRef\]](#) [\[PubMed\]](#)
176. Heidemann, A.; Hülter, S.M.; Tekieli, M. Machine Learning with Real-World HR Data: Mitigating the Trade-off between Predictive Performance and Transparency. *Int. J. Hum. Resour. Manag.* **2024**, *35*, 2343–2366. [\[CrossRef\]](#)
177. Lera-Leri, R.X.; Liscio, E.; Bistaffa, F.; Jonker, C.M.; Lopez-Sanchez, M.; Murukannaiah, P.K.; Rodriguez-Aguilar, J.A.; Salas-Molina, F. Aggregating Value Systems for Decision Support. *Knowl.-Based Syst.* **2024**, *287*, 111453. [\[CrossRef\]](#)
178. Şentürk, Ö. AI-Driven Climate Adaptation: Technical Applications, Ethical Governance, and Social Inclusion. *Acadlore Trans. Mach. Learn.* **2026**, *5*, 20–31. [\[CrossRef\]](#)
179. Tuyen, V.V. Human Behavioral Dynamics in AI-Assisted Decision Making: An Integrated SWOT-AHP-TOPSIS Analysis. *J. Intell. Manag. Decis.* **2026**, *5*, 64–76. [\[CrossRef\]](#)
180. Dong, X.H.; Zhu, X.Q.; Li, L.P. Digital Finance and Industrial Chain Resilience in China: A Spatial Network Perspective. *J. Intell. Manag. Decis.* **2026**, *5*, 47–63. [\[CrossRef\]](#)
181. Inder, S.; Sood, K.; Grima, S. Antecedents of Behavioural Intention to Adopt Internet Banking Using Structural Equation Modelling. *J. Risk Financ. Manag.* **2022**, *15*, 157. [\[CrossRef\]](#)
182. Awuraris, A.K.; Daniel, R.; Muthukumar, P. Empowering Accessibility to Digital Space Through Generative AI to Support People with Disabilities. *Acadlore Trans. Mach. Learn.* **2026**, *5*, 44–57. [\[CrossRef\]](#)
183. Oppong-Fosu, K.; Obagbuwa, O. Financial Inclusion Driven by Digital Financial Platforms: Impact on Economic Growth in Ghana. *J. Account. Financ. Audit. Stud.* **2025**, *11*, 244–258. [\[CrossRef\]](#)
184. Akbulut, O.Y. Assessing the Environmental Sustainability Performance of the Banking Sector: A Novel Integrated Grey Multi-Criteria Decision-Making (MCDM) Approach. *Int. J. Knowl. Innov. Stud.* **2024**, *2*, 239–258. [\[CrossRef\]](#)
185. Varma, P.; Nijjer, S.; Sood, K.; Grima, S.; Rupeika-Apoga, R. Thematic Analysis of Financial Technology (Fintech) Influence on the Banking Industry. *Risks* **2022**, *10*, 186. [\[CrossRef\]](#)

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