International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group | Volume 7 Issue 1 PP 322-331, 2026
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V7I1P139

Original Article

Trustworthy AI in Financial Risk Management: Applications for SME Compliance, Consumer Protection, and Audit Readiness

Norman Alfredo Gamboa Gamarra
Attorney (Colombia-licensed) & Enrolled Agent IRS certified. Financial Risk & Tax Compliance Consultant, UNiversidad Autonoma de Bucaramanga and CEO Momentum Factory LLC.

Received On: 04/02/2026 ***Revised On: 05/03/2026*** ***Accepted On: 08/03/2026*** ***Published On: 12/03/2026***

Abstract *- The digital revolution that is taking place rapidly in the financial sector brings about the incorporation of AI in risk management, compliance, and decision-making. As financial institutions and other small and medium-sized enterprises (SMEs) increasingly depend on AI to assess creditworthiness, detect fraud, report to regulators, and manage portfolios, the question is whether such systems can be trusted. This paper will explain the origins, usage, and regulation of Trustworthy AI in Financial Risk Management. It describes the main principles, including fairness, transparency, accountability, privacy, security and robustness and addresses the issue of bias, explainability, regulatory alignment, and model risk. The paper also addresses AI-based compliance solutions as used by SMEs and their contribution to efficiency enhancement, enhanced efficacy with regulations, and audit preparedness. It also emphasizes ethical governance, consumer protection and AI assurance systems required to keep the finances stable and to trust the people. The paper offers an automated approach to simulating responsible and sustainable AI systems in dynamic financial environments by incorporating the lifecycle management, documentation standards and perpetual validation or practices. The findings demonstrate the need to strike the appropriate balance between innovation and regulation, ethical protection. Finally, the study will contribute to the development of AI-based financial risk management solutions that are resilient, open, and compliant.*

Keywords *- Trustworthy AI, SME Compliance, Ethical Governance, Consumer Protection, Regulatory Compliance, Bias Detection.*

1. Introduction

The digital transformation has placed the financial industry under a lot of pressure, because more people are willing to apply the use of artificial intelligence (AI) to assess credit, detect frauds and optimize the portfolio, and to comply with regulations. Alongside this is the growing regulatory attention, risk landscape, and growth of transparency and accountability demands on financial institutions and SMEs [1]. Although AI-based systems might considerably improve the effectiveness and predictability, the issue of model bias, explainability, robustness, and regulatory alignment has led to the need to create trustworthy AI frameworks. The credibility of AI systems is very important in Financial Risk Management where the decisions of the systems directly impact on the allocation of capital, consumer rights, and institutional stability.

The principles of fairness, transparency, accountability, privacy, robustness, and adherence to regulations are included in trustworthy AI in financial risk management [2]. In the case of SMEs that in most instances have limited technical and compliance capacity, AI-based risk management software should not only be capable of achieving operational efficiency, but also compliance with complex regulatory provisions [3]. The combination of explainable AI (XAI) applications, pipeline of auditable models, and governance frameworks allows SMEs to prove compliance and remain competitive. In addition, effective consideration of ethical and legal concerns in the design of AI systems helps to adopt the system sustainably and resolve reputational and legal risks.

The other important aspect of responsible AI implementation is consumer protection. The unconscious transmission of systemic biases in algorithmic decision-making in lending, insurance underwriting, credit scoring, and various other algorithms can result in opaque outcomes that disadvantage vulnerable groups [4][5]. By embracing fair predictive methods, detection of bias and model documentation, the financial institutions can safeguard the rights of the consumers and maintain predictive accuracy. In this regard, trustworthy AI is a technological protection, as well as an instrument for increasing the trust of people in the online financial setting.

Audit readiness also helps improve the significance of quality AI frameworks in financial risk management. Data sources, model assumptions, validation procedures and decision logic are needed by regulatory bodies. AI-based technologies and systems are meant to facilitate internal audits and regulatory checks because they might include some capabilities of built-in monitoring, version control, and documentation [6]. These convergences of plausible AI principles and financial risk management practices can in these ways offer a prioritized route by which SMEs and financial institutions may find a way of complying with the regulations, advance consumer protection and increase organizational resilience to the changing regulatory environment.

1.1. Structure of the Paper

There are six key sections in this paper. In Section II, the Foundations of Trustworthy AI in Financial Risk Management are discussed. Section III deals with Applications to SME Compliance. In financial risk management, Section IV discusses the Ethical Governance, Consumer Protection, and AI Assurance. Section V shows the Literature Review. Lastly, there is Section VI, which gives the Conclusion and Future Work.

2. Foundations Of Trustworthy Ai In Financial Risk Management

Due to the fast development of AI technologies across a range of sectors, policymakers, developers, and researchers are now prioritising the credibility of AI [7]. The following characteristics are indicative of trustworthy AI: validity, dependability, transparency, accountability, privacy, and fairness in system design. Acquiring trust requires integrating user requirements, validating the system, and continuously monitoring performance. Fairness, explainability, privacy, accountability, and openness are all aspects of trustworthiness that should be carefully considered in such endeavours, with each dimension being balanced, appropriately prioritised, and optimised according to the specific circumstances [8]. AI systems that are founded on these principles are capable of mitigating societal risks and ensuring that AI usage is consistent with widely held human values.

2.1. Components of AI Trustworthiness

Credibility in AI has a number of major elements that form the credibility and moral correctness of AI systems. Figure 1 provides a summary of the following points: security, privacy, openness, accountability, and impartiality.

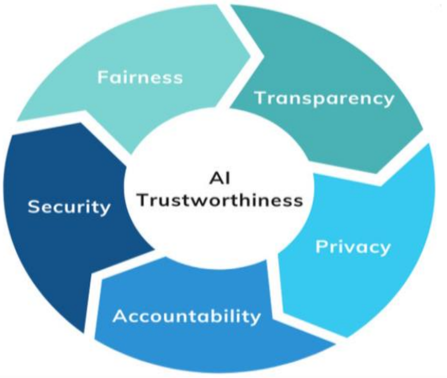

Fig 1: Elements of AI Reliability: Privacy, Accountability, Fairness, Transparency, And Security

2.1.1. Fairness

Reliable AI is predicated on the principle of fairness, which provides AI systems with reduced damage or the absence of biases, as AI systems may produce discriminatory results. People's trust, human autonomy, and discrimination and inequality are the biases associated with ethical issues [9]. Decisions influenced by AI may have a profound effect on people's lives, making it especially pertinent in delicate fields like healthcare, criminal justice, and finance.

2.1.2. Transparency

Transparency is a fundamental aspect of trusts and accountability, so in the case of AI, transparency must exist by revealing and exposing decision-making to enable its users to comprehend the process of decision-making used to make predictions [10]. This is of particular concern in black-box models. LIME and SHAP are the techniques that are said to promote transparency.

2.1.3. Privacy

AI privacy concerns need to protect the data of users against unauthorized access and to guarantee the safe usage of personal information [11], especially in such sensitive fields as healthcare and finances. AI privacy protection from the outset necessitates techniques such as Differential Privacy (DP), which introduces statistical noise to datasets to protect individual identities, and Federated Learning (FL), which allows models to be trained on distant devices without sending raw data.

2.1.4. Accountability

Ethical AI governance requires accountability, which means that systems should be run in a manner within ethical systems that are transparent and regulatory [12]. Accountability, when applied to trustworthy AI, means that the responsibility of decisions made by AI systems is explicitly assigned, so that any stakeholder can be able to trace any action to a particular individual or object.

2.1.5. Security

AI security guarantees that AI systems are not vulnerable to attacks, their integrity, confidentiality and privacy remain safe, particularly when subjected to malicious attacks [13]. There are two ways to attack AI systems: model extraction, where the attacker queries the system to try to rebuild or steal the model, and evasion, when the attacker produces inputs designed to make the AI system forecast incorrectly.

2.2. Importance Across Sectors

The need for reliable AI cuts across various industries, with each one having its specific needs and issues. The most intelligent people work in the most important sectors, including healthcare, banking, government, and autonomous systems, where trustworthy AI is essential for building confidence in the AI technologies listed in Table I.

Table 1: Sector-Specific Applications and Trustworthiness Considerations of AI

Sector	Key AI Applications	Trustworthiness Requirements	Governance / Frameworks / Initiatives
Finance[14]	Fraud detection, credit scoring, algorithmic trading, risk prediction, regulatory compliance automation	Fairness, transparency, accountability, security, regulatory compliance	National AI strategies (e.g., Singapore, Denmark, USA), Frankfurt Institute for Risk Management and Regulation (FIRM), financial regulatory frameworks emphasizing ethical AI
Healthcare [15]	Diagnostics, predictive modeling, treatment planning, patient monitoring, virtual health assistants, wearable technologies	Interpretability (LIME, SHAP), fairness, privacy protection, reliability, robust performance evaluation	Privacy-Preserving Machine Learning (PPML), differential privacy, federated learning, clinical evaluation metrics (Precision, Recall, F1, MCC)
Public Administration[16][17]	Public service delivery, criminal justice analytics, social welfare management, resource allocation	Transparency, accountability, fairness, citizen trust, responsible data governance	eGovernment initiatives, inter-agency collaboration frameworks, digital governance strategies
Autonomous Driving & Robotics [18]	Autonomous navigation, robotics control systems, sensor-based decision-making	Safety, reliability, robustness, ethical reasoning, cybersecurity resilience	Asilomar Principles, Partnership on AI (PAI), IEEE Ethically Aligned Design, IEEE 7000-series standards, British Standard BS 8611:2016

2.3. Challenges in Achieving Trustworthy AI

Although AI has been used by many industries and has an influence on the business, making AI reliable is a significant issue. The decision-making within AI systems is usually hidden due to the complexity of such systems, especially the fact that the ML models are often black-box, which is harmful to transparency and accountability. Furthermore, privacy concerns, trade-offs between competing trustworthiness measures, and biases in training data make it increasingly challenging to develop morally and reliably sound AI systems. The issues are only exacerbated by the fact that new threats, including adversarial offensive, misinformation, and data poisoning, are becoming a problem and are abusing the flaws in AI technologies. To overcome these problems, there should be interdisciplinary efforts to create strong technical protection, ethical governance models [19], and continuous research to find a balance between the reliability, security, and the influence of AI. Promotions done by the society are also necessary so that the development of AI does not violate the expectations of society [20]. The current paper discusses the main obstacles to establishing trustworthy AI, highlighting the risks involved and the current efforts to deal with them. Figure 2 provides a summary of the difficulties.

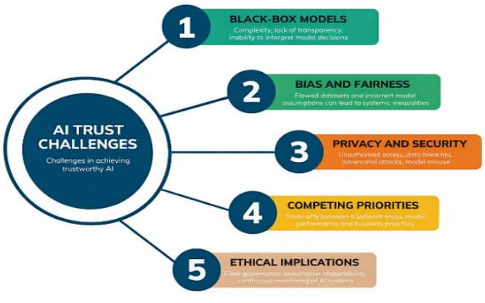
AI TRUST CHALLENGES
Challenges in achieving trustworthy AI
1
BLACK-BOX MODELS
Complexity, lack of transparency, inability to interpret model decisions
2
BIAS AND FAIRNESS
Flawed datasets and incorrect model assumptions can lead to systemic inequalities
3
PRIVACY AND SECURITY
Unauthorized access, data breaches, adversarial attacks, model misuse
4
COMPETING PRIORITIES
Trade-offs between trustworthiness, model performance and business priorities
5
ETHICAL IMPLICATIONS
Clear governance, stakeholder responsibility, continuous monitoring of AI systems

Fig 2: Key Challenges in Achieving Trustworthy AI

Their "black-box" nature refers to the opaqueness of AI and ML models' decision-making processes, especially those used in DL models. Such models are also not very transparent and interpretable, and it is difficult to comprehend how the decisions are made by the users [21].

2.3.1. Bias and Fairness

Prejudice is an aspect that plagues AI systems, which influences their impartiality, trustworthiness, and moral soundness. It might occur throughout any phase of the AI lifecycle, such as problem formulation, data collection, model training, and deployment.

2.3.2. Privacy and Security

The emergence of data-intensive applications that run on AI has resulted in major developments but has also introduced the issue of privacy and security. These issues are encountered in areas like medical services, finances, and personalized marketing, among which sensitive information is often utilized to train and implement AI models.

2.3.3. Balancing Competing Trust Components in AI Systems

To ensure trust in AIsystems, one must pay special attention to the trade-offs among the most prominent aspects of the field like privacy, transparency, fairness, and efficiency [22]. These qualities tend to pose conflicting requirements and it is difficult to work with them both at the same time.

2.3.4. Accountability and Ethical Implications

The process of ensuring accountability in AI systems is not a simple task because of the complexity of contemporary ML systems. Stakeholders' capacity to justify AI system actions and choices in a way that satisfies cultural norms and ethical standards is one definition of accountability [12].

2.3.5. Interactions and Trade-Offs Between Trustworthiness Metrics

Accuracy, robustness, explainability, fairness, and privacy are some of the most popular metrics for evaluating AI reliability, these metrics do not always evolve in a consensus-building process. When put into practice, improvements in one dimension may come at the expense of other dimensions' performance. It is important to understand and examine these tensions to be capable of creating a factual portrait of what it will appear to trust AI in the real world.

2.4. Risk Governance and AI Lifecycle Management

It is essential that the transparency, reliability, and compliance of AI-based financial risk systems are central to risk governance and AI lifecycle management [23]. Where risk governance forms the governance systems and structures of accountability and ethical controls, lifecycle management triggers a systematic production, validation, control and constant enhancement of AI designs. Table II gives a summary of the key dimensions of the two components.

Table 2: Risk Governance and AI Lifecycle Management Dimensions in Financial Risk AI

Risk Governance Dimension	AI Lifecycle Management Dimension
Regulatory-aligned AI governance frameworks	Structured lifecycle stages (data → development → validation → deployment → monitoring → retirement)
Defined roles and accountability (developers, risk managers, compliance, senior management)	Data quality management and governance controls
Model Risk Management (validation, approval, monitoring, decommissioning)	Integration of Explainable AI (XAI) techniques
Ethical AI integration (fairness, transparency, accountability)	Rigorous model validation and performance evaluation
Risk assessment protocols (bias, explainability, cybersecurity, resilience)	Continuous monitoring and drift detection
Documentation, audit trails, and independent oversight	Change management and version control
Stress testing and scenario analysis	Privacy-preserving techniques and incident response mechanisms
Cross-functional collaboration and regulatory compliance	Periodic lifecycle audits and compliance reviews

3. Applications For SME Compliance

Reliable AI would enhance the compliance of SMEs in financial risk management by enhancing efficiency, transparency, and regulatory alignment. It allows SMEs to transition between reactive compliance strategies and proactive and data-driven risk management systems [24]. AI enhances regulatory trust and trust among the stakeholders by integrating the concepts of explainability, fairness and accountability. Key applications include:

3.1. Automated Regulatory Monitoring and Reporting

AI systems scan regulatory change, connect that to internal processes, and produce precise compliance reports. Explainable models make transparency and audit-ready and decrease manual compliance burdens [25]. This enhances consistency in reporting as well as enabling SMEs to act promptly on the changes in their regulatory needs.

3.2. Anti-Money Laundering (AML) and Fraud Detection

AI identifies suspicious activities and transactions in real time. Explainable and fair systems minimize false positives, enhance internal controls and aid regulatory compliance [26]. The capacity of continuous learning also improves the detection of new risks of financial crimes by the system.

3.3. Credit Risk Assessment and Responsible Lending

Artificial intelligence-based credit models enhance the assessment of risks based on financial and other data. Fair, consistent and compliant lending decisions are guaranteed through transparency and prevention of bias. This improves the quality of the portfolio and reduces both legal and reputational risks.

3.4. Data Protection and Privacy Compliance

AI helps in classification of data, detecting breaches, and cybersecurity monitoring. Strong and safe systems can assist SMEs in adhering to the data protection legislation and reduce legal and reputational threats. Data governance and data incident response mechanisms are further enhanced by automated monitoring [27].

3.5. ESG and Sustainability Reporting

AI automatizes the ESG data collection and reporting processes, which provide traceable and reliable reports that address regulatory and investor expectations and make sustainable finance more accessible [28]. It also enhances accuracy and comparability of sustainability measures.

3.6. Internal Risk Governance and Audit Support

AI keeps track of critical risk factors, identifies compliance violations and produces audit trails [29]. This enhances the governance structures and increases the readiness of auditing. Predictive analytics also assists in proactive risk management and planning.

3.6.1. Importance of Risk Compliance

Risk compliance is defined as the commitment of a financial institution to comply with regulatory requirements and legal provisions in a bid to conduct a lawful business and ensure stability of business. Intense regulatory control has been critical in the management of liquidity risk, minimizing financial stress, eliminating fraud and accountability especially following the global financial crisis, which brought tougher regulations on big institutions. Compliance is not just a legal requirement, but an effective strategic tool, which strengthens the competitive advantage and shareholders' value [30].

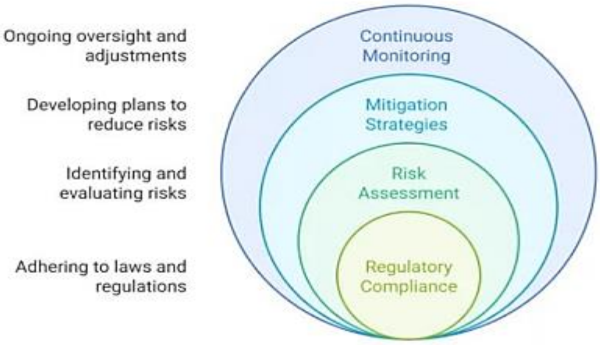
Ongoing oversight and adjustments
Continuous Monitoring
Developing plans to reduce risks
Mitigation Strategies
Identifying and evaluating risks
Risk Assessment
Adhering to laws and regulations
Regulatory Compliance

Fig 3: Compliance and Risk Management

Nonetheless, the risk of compliance has increased tremendously and regulators are subjecting the systems and standards of financial institutions to high levels of scrutiny. Consequently, companies are putting a lot of money into well-developed compliance programs to avoid breaches and streamline controls (as shown in Figure 3). The difficulty of having any effective balance is that the institutions have to focus on risk, complexity and cost-based regulations without leaving out low-risk regulations and without making their compliance costs too high.

3.6.2. Regulatory Frameworks

- The aftereffects of previous crises, such as the global financial crisis, are still being felt and are working against the consolidation of global financial regulation.
- The swift evolution of AI-financial products has far exceeded the current regulatory provisions, leaving regulatory loopholes.
- In others, the new rules have correspondingly slowed the rate at which banks can compete and adjust to market demand, which brings about the issue of systemic instability and cascading disruption of an interconnected global economy.
- The increasing power of opaque sovereign wealth funds has made public finance decisions difficult for governments [31], and it usually gives rise to political and social tensions.
- Open banking and digital financial platforms have enhanced consumer choice but have also intensified misinformation risks, politicization risks, and financial instability risks in weak states.
- Without the human common sense, AI systems represent a delicate balance between positive and disastrous consequences and thus, a proper regulation is a crucial requirement [32].
- In spite of the banks reporting AI systems to the national regulators, innovation can rise beyond the existing controls.
- The most common responses to AI-related risks by regulators are restricting their conduct to high standards or promoting the use of standardized models in the industry.
- This creates a perpetual race between innovation and control, particularly in the presence of no global frameworks and perpetually implemented AI governance policies.
- Too much regulatory restrictiveness can reduce creative financial innovation, limit the development of superior models, and enhance system vulnerability [33].
- Finally, given the lack of balance and adaptive control, the loss of informational benefit may take place among the public organizations, and the companies can be characterized by limited expansion because of limited incorporation of sophisticated financial skills.

3.6.3. Key Principles for Governance Tools for SMEs

- Affordability: Offer affordable solutions that can be afforded by the small and medium-sized businesses [34]. Case Study: SME-specific low-cost subscription models of AI auditing software.
- Ease of Use: Make the tools of governance simple so that they can easily be adopted and implemented by non-technical users.
- Scalability: Solutions to design that allow the company to scale with its AI requirements.
- Regulatory Alignment: Make sure that the tools are in accordance with the applicable regulations including GDPR and that they can ease the legal burden.
- Customization: Modify tools according to needs and industries of SMEs.

4. Ethical Governance, Consumer Protection, and AI Assurance in Financial Risk Management

With the fast-paced nature of the financial environment, the introduction of AI to the risk management procedures has introduced both immense opportunities and ethical challenges of considerable magnitude. Ethical governance, consumer protection, and AI assurance have thus become important elements in Trustworthy AI in financial risk management [35]. Since institutions are implementing AI in credit assessment, fraud detection, underwriting and compliance monitoring, it is essential that systems be fair, transparent, secure and accountable in their operation [36]. This requires that adequate bias detection exists, that automated decision making is adequately explained, that the privacy of data exists, that ongoing model validation is undertaken and audit mechanisms well documented. The implementation of this principle in AI models will introduce regulatory compliance, as well as consumer confidence and enable sustainable financial innovations.

4.1. Bias Detection and Fair Lending Practices

Artificial intelligence systems that have lending options and risk analysis should be highly secured to prevent unfair and discriminatory results. Algorithms rely on the past, and thus there is the probability of the automated judgments being based on the existing biases which can be social and economic biases [37]. It is advised that banks ought to frequently examine the fairness of their operations and make corrective action in the event of any imbalance of their operations. To ensure that all the applicants are treated equally, human control and compliance with the fair lending standards should be enforced.

4.2. Transparency in Credit and Insurance Decision Making

The consumers must be in a position to learn how credit approvals, lending decisions with regard to pricing of loans or underwriting decisions on insurance are made. The AI models should then be in a position to provide clear and understandable accounts as compared to black boxes. Accountability will increase transparency in decision-making, conflict reduction and customer trust. It will also help to satisfy regulatory mandates that require the explanation of automated financial decisions.

4.3. Data Privacy and Security Safeguards

Financial systems guided by AI require massive amounts of sensitive information, which has made data protection a high-priority issue [38]. The institutions should have safe data collection, storage and processing. Cybersecurity, encryption, and access control systems are highly developed and prevent unauthorized use or breaking. The privacy-by-design concepts will be implemented in the AI systems and will help to guarantee compliance and customer trust.

4.4. Model Documentation and Traceability

Documentation makes artificial intelligence models responsibility and readable. Institutions are to document model design procedures, data sources, and the output of the process of validation, and performance measures. Traceability enables organizations to restructure the way in which some of their decisions were made that is crucial especially in the audit or regulatory audit. Table III shows key components of Model Documentation and Traceability below:

Table 3: Key Components of Model Documentation and Traceability

Component	Purpose	Benefit for Governance
Model Design Documentation	Records algorithms, assumptions, and methodologies used	Enhances transparency and accountability
Data Source Records	Identifies training and testing datasets	Supports data integrity and compliance checks
Validation Reports	Documents performance testing and stress scenarios	Ensures reliability and regulatory alignment
Version Control Logs	Tracks updates and modifications to models	Facilitates audit readiness and change management
Decision Audit Trails	Captures model outputs and decision logic	Enables traceability during inspections

4.5. Continuous Validation and Performance Monitoring

The use of AI systems should be regularly evaluated in order to achieve consistency in accuracy and fairness. The changes in the market conditions or data patterns can influence the model's performance over time. The reliability of the system includes constant verification by recalibration and review by independent persons who minimise operational risks.

4.6. Regulatory Inspection and Audit Support Mechanisms

Banking organizations are also supposed to provide the formal systems that can support the regulatory audits and internal audits of AI systems. It must also have audit trails, governance policy and well-defined accountability structures that will help them manage supervisory review [39]. The presence of effective oversight mechanisms shows responsible AI use, as well as provides greater regulatory trust in financial risk management practices.

5. Literature Review

The current study indicates that AI is increasingly being included in real-time and predictive systems for financial risk management and compliance. Nevertheless, certain problems like regulatory loopholes, transparency, and ethics continue to be major challenges. Table IV gives a comparative summary.

Nastoska et al. (2025). In this research, a comparative study of the existing standards is done, and the application of the standards in various industries is demonstrated to indicate their effectiveness. Case studies in the real world in the application of trust metrics to healthcare, financial services, and autonomous systems applications are shown by real-world case studies. It is found that establishing trustworthiness means finding trade-offs between conflicting measures, including just versus efficient or secret versus transparent, and the significance of the interdisciplinary cooperation as a source of strong AI governance. The new trends imply the necessity to develop adaptive frameworks of AI trustworthiness that would be changed along with improvements in AI technologies [40].

Vyas (2025) examines how AI is being used in multinational financial institutions' risk setup, focusing on risk rating, anomaly detection, predictive analytics, and real-time decision-making. Through consideration of the existing use cases in organizations such as JPMorgan Chase and PayPal, we consider how AI has transformed practices and policy structures in the United States, Europe, and in developing economies in Asia and Africa. The analysis of the qualitative data of the case studies and white papers is combined with the quantitative data analysis using the public financial data, trends in the adoption of AI, and predictive models. We also predict the future of AI in risk management in the five years to come, where it will make a difference in reducing the financial losses, enhancing compliance and streamlining the underwriting processes [41].

Xie, Tan and Liu (2024) provide an overview of the concepts, kinds, and constraints of traditional methods for managing financial risk. It then looks at the use of AI, namely in the detection, evaluation, monitoring, and reporting of credit risk, as well as how AI may help create more sophisticated risk prediction models and its main role in supporting real-time risk management decision-making. Suggesting the future of the integration of AI and blockchain technology into managing financial risk, including automating the execution of smart contracts, enhancing the transparency of the supply chain, and exploring the potential for financial inclusiveness, the prospects are impressive. By studying, the objective is to offer and give a holistic view to financial institutions and other concerned researchers to have a better grasp and use AI innovative practices in financial risk management [42].

El Hajj and Hammoud (2023) use a mixed-method research approach integrating quantitative surveys with qualitative literature reviews to investigate the potential applications of AI and ML in the financial markets. There is quantitative evidence that the banking sector is heavily using AI and ML. Risk management, algorithmic trading, CreditScoring, fraud detection, and customer services are some of the most prevalent uses of this technology. The qualitative research also shows the main points, which include the following: the role of regulation, the change of the workforce, ethical and societal concerns, and the trends in the use of AI and ML [43].

S Anamol, A Anjaneyulu, and M Raja Aseem (2023) focus on how AI may be used to forecast and prevent a range of financial hazards, such as market, operational, and credit risks. Financial institutions may utilise ML algorithms, NLP, and sophisticated data analytics to enhance their risk assessment methods, fraud detection abilities, and regulatory compliance. In addition to enhancing accuracy and efficiency, the results show that AI-based risk management systems provide real-time data that may be used for proactive decision-making. The impact of AI on building a stronger and safer financial system is shown in this research [44].

Fritz-Morgenthal, Hein and Papenbrock (2022b) highlight the use of existing methods, tools, and platforms to assist in the creation of reliable, explicable, auditable, and managed AI/ML for production. It also endeavours to provide practical recommendations for establishing a risk-based testing and governance framework for these models. This paper is valuable because it encourages further thought outside of the financial services industry, particularly in relation to what the aforementioned EU consultation referred to as High Risk models, which is relevant to the recent publication of the European AIA by the European Union [45]

Table 4: Comparative Analysis of ML/DL-Based Intrusion Detection Studies

Authors & Year	Primary Focus	Key Challenges	Major AI Applications	Key Contributions	Future Implications
Nastoska et al. (2025)	AI trustworthiness standards and cross-industry governance	Trade-offs between fairness vs. efficiency and privacy vs. transparency; need for interdisciplinary coordination	Trust metrics in healthcare, financial services, and autonomous systems	Demonstrates that AI trust requires balancing competing ethical metrics; highlights role of adaptive governance	Calls for evolving, adaptive AI trust frameworks aligned with technological advancements
Vyas (2025)	Integration of AI into global financial risk infrastructure	Regulatory differences across regions; data governance issues; managing predictive model risks	Predictive analytics, anomaly detection, real-time decision-making, risk scoring	Shows AI reshaping risk practices in institutions like JPMorgan Chase and PayPal; combines qualitative and quantitative insights	The anticipated involvement of AI in the refinement of underwriting, the enhancement of compliance, and the reduction of financial losses
Xie, Tan & Liu (2024)	AI applications in financial risk identification, assessment, and monitoring	Limitations of traditional methods; integration complexity; regulatory lag	Risk prediction models, real-time decision support, AI-blockchain integration	Highlights AI's role in intelligent risk prediction and smart contract automation	Emphasizes blockchain integration for transparency, automation, and financial inclusion
El Hajj & Hammoud (2023)	Adoption and impact of AI/ML in financial markets	Regulatory uncertainty; ethical concerns; workforce transformation; barriers to adoption	Algorithmic trading, fraud detection, credit scoring, risk management, customer service	Confirms growing AI/ML adoption; identifies governance and ethical challenges	Stresses importance of regulation, workforce upskilling, and ethical oversight
S. Anamol, A. Anjaneyulu &	AI for predicting and mitigating	Data quality limitations;	Credit risk, market risk,	AI improves accuracy, efficiency,	Supports AI's transformative

M. Raja Aseem (2023)	financial risks	interpretability concerns; cybersecurity and compliance risks	operational risk, fraud detection	and real-time decision-making capabilities	potential in building resilient financial ecosystems
Fritz-Morgenthal, Hein & Papenbrock (2022b)	Governance and testing frameworks for trustworthy AI/ML	Managing "high-risk" AI systems; ensuring explainability, auditability, and compliance with evolving regulations	Risk-based governance frameworks, model validation, explainable AI	Provides practical guidance for responsible AI implementation in production	Aligns governance models with the EU AIA and high-risk AI oversight

6. Conclusion and Future Work

The growing use of artificial intelligence in financial systems is a paradigm shift in risk identification, assessment and management. The paper has put the emphasis on the need to introduce the plausible AI concepts that include fairness, transparency, accountability, privacy, resiliency and regulation in the financial risk management systems. The paper will show that it is not only that reliable AI should be demanded so that its functioning is efficient, but also it will save consumers and keep the institutions stable. Risk governance, lifecycle management and continuous monitoring play important roles in reducing the bias, model drift, cybersecurity threats and compliance failures. Future studies would be oriented to constructive regulatory models that would adapt to the emerging AI innovations especially in the fields of explainable deep learning, real-time supervisory technologies and the integration of AI blockchain. There is also the necessity to be more concerned with the cross-border alignment of regulations and AI standards in the specific sector. Furthermore, further empirical studies of the evaluation of the long-term operations, fairness, and resilience of AI-based risk systems would also be beneficial in evidence-based policymaking and responsible financial innovation.

References

[1] V. Verma, "Security Compliance and Risk Management in AI-Driven Financial Transactions," *Int. J. Eng. Sci. Math.*, vol. 12, no. 7, pp. 107–121, 2023.

[2] S. Kumara, "Identity-Driven IoT Security in Telecom Ecosystems: Implications for Scalable and Trustworthy Digital Infrastructure," *Int. J. Appl. Math.*, vol. 38, no. 12s, pp. 2797–2816, Dec. 2025, doi: 10.12732/ijam.v38i12s.1588.

[3] C. Gardner, K. M. Robinson, C. J. Smith, and A. Steiner, "Contextualizing end-user needs: How to measure the trustworthiness of an AI system," *Carnegie Mellon Univ. Softw. Eng. Inst.*, 2023, doi: 10.1184/R1/23819100.

[4] N. Khalid, A. Qayyum, M. Bilal, A. Al-Fuqaha, and J. Qadir, "Privacy-preserving artificial intelligence in healthcare: Techniques and applications," *Comput. Biol. Med.*, vol. 158, p. 106848, May 2023, doi: 10.1016/j.compbiomed.2023.106848.

[5] V. Shah, "Managing Security and Privacy in Cloud Frameworks : A Risk with Compliance Perspective for Enterprises," *Int. J. Curr. Eng. Technol.*, vol. 12, no. 6, pp. 606–618, 2022, doi: 10.14741/ijcet/v.12.6.16.

[6] U. K. Deloitte, "Banking on the bots: unintended bias in AI," *Deloitte London, UK*, 2023.

[7] C. Cousineau, R. Dara, and A. Chowdhury, "Trustworthy AI: AI developers' lens to implementation challenges and opportunities," *Data Inf. Manag.*, vol. 9, no. 2, p. 100082, Jun. 2025, doi: 10.1016/j.dim.2024.100082.

[8] B. Li *et al.*, "Trustworthy AI: From Principles to Practices," *ACM Comput. Surv.*, vol. 55, no. 9, pp. 1–46, Sep. 2023, doi: 10.1145/3555803.

[9] E. Ferrara, "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies," *Sci*, vol. 6, no. 1, p. 3, Dec. 2023, doi: 10.3390/sci6010003.

[10] H. Q. Ngo, T. H. Nguyen, and G. Du Kang, "The effect of perceived value on customer engagement with the moderating role of brand image: A case study in Vietnamese restaurants," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 7C2, pp. 451–461, 2019.

[11] S. H. Deep Kolagani, M. Bhandar, and R. Altounjy, "Enhancing DevOps Security with LLMs for Automation and Compliance," in *2025 7th International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, IEEE, Oct. 2025, pp. 1514–1519. doi: 10.1109/ICIDCA66325.2025.11280531.

[12] C. Novelli, M. Taddeo, and L. Floridi, "Accountability in artificial intelligence: what it is and how it works," *AI Soc.*, vol. 39, no. 4, pp. 1871–1882, Aug. 2024, doi: 10.1007/s00146-023-01635-y.

[13] A. Oseni, N. Moustafa, H. Janicke, P. Liu, Z. Tari, and A. Vasilakos, "Security and Privacy for Artificial Intelligence: Opportunities and Challenges," *arXiv*, Feb. 2021.

[14] K. Ashraf, S. Nawar, M. H. Hosen, M. T. Islam, and M. N. Uddin, "Beyond the Black Box: Employing LIME and SHAP for Transparent Health Predictions with Machine Learning Models," in *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS)*, IEEE, Mar. 2024, pp. 1–6. doi: 10.1109/iCACCESS61735.2024.10499522.

[15] S. Fritz-Morgenthal, B. Hein, and J. Papenbrock, "Financial risk management and explainable, trustworthy, responsible AI," *Front. Artif. Intell.*, vol. 5, p. 779799, 2022.

[16] C. van Noordt and L. Tangi, "The dynamics of AI capability and its influence on public value creation of

AI within public administration," *Gov. Inf. Q.*, vol. 40, no. 4, p. 101860, oct. 2023, doi: 10.1016/j.giq.2023.101860.

[17] M. Kuziemski and G. Misuraca, "AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings," *Telecomm. Policy*, vol. 44, no. 6, p. 101976, Jul. 2020, doi: 10.1016/j.telpol.2020.101976.

[18] K. Michael, R. Abbas, G. Roussos, E. Scornavacca, and S. Fosso-Wamba, "Ethics in AI and autonomous system applications design," *IEEE Trans. Technol. Soc.*, vol. 1, no. 3, pp. 114–127, 2020, doi: 0.1109/TTS.2020.3019595.

[19] S. Singamsetty, "Efficacy of Data Governance: A Cutting Edge Approach to Ensuring Data Quality in Machine Learning for Banking Industry," in *2024 2nd International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, IEEE, Dec. 2024, pp. 1–7. doi: 10.1109/SCOPES64467.2024.10991944.

[20] N. Polemi, I. Praça, K. Kioskli, and A. Bécue, "Challenges and efforts in managing AI trustworthiness risks: a state of knowledge," *Front. Big Data*, vol. 7, p. 1381163, May 2024, doi: 10.3389/fdata. 2024.1381163.

[21] D. Kaur, S. Uslu, K. J. Rittichier, and A. Durresi, "Trustworthy Artificial Intelligence: A Review," *ACM Comput. Surv.*, vol. 55, no. 2, pp. 1–38, Feb. 2023, doi: 10.1145/3491209.

[22] S. Narang and V. Gopi Kolla, "Next-Generation Cloud Security: A Review of the Constraints and Strategies in Serverless Computing," *Int. J. Res. Anal. Rev.*, vol. 12, no. 3, 2025, doi: 10.56975/ijrar.v12i3.319048.

[23] N. K. R. Choppa and J. W. Sajja, "Towards Human-in-the-Loop Orchestration of Agentic SAP Ecosystems: Governance, Cognitive Impact, and Control Tower Design," *J. Inf. Syst. Eng. Manag.*, vol. 10, 2025.

[24] H. A. Ahmed, A. Hameed, and N. Z. Bawany, "A Survey on Regulatory Compliance and AI-Based Risk Management in Financial Services," *Int. J. Artif. Intell. Data Sci. Mach. Learn.*, vol. 4, p. e820, Dec. 2023, doi: 10.63282/3050-9262.IJAIDSML-V4I4P106.

[25] E. Leonard, B. Sheehan, M. Mullins, and D. Shannon, "Generative AI for Enhanced Risk Management in SMEs Date: 3rd July 2024," 2025. doi: 10.2139/ssrn.5213414.

[26] K.-C. Yao *et al.*, "Application of Generative AI in Financial Risk Prediction: Enhancing Model Accuracy and Interpretability," *Information*, vol. 16, no. 10, p. 857, oct. 2025, doi: 10.3390/info16100857.

[27] S. Singamsetty, "AI-Based Data Governance: Empowering Trust and Compliance in Complex Data Ecosystems," *Int. J. Comput. Math. Ideas*, vol. 13, no. 03, pp. 1007–1017, 2021, doi: 10.70153/IJCMI/2021.13301.

[28] J. W. Sajja, G. B. Komarina, and N. K. R. Choppa, "The Convergence of Financial Efficiency and Sustainability in Enterprise Cloud Management," *J. Comput. Sci. Technol. Stud.*, vol. 7, no. 4, May, pp. 964–992, 2025, doi: 10.32996/jcsts.2025.7.4.110.

[29] S. Paleti, "The Role of Artificial Intelligence in Strengthening Risk Compliance and Driving Financial Innovation in Banking," *Int. J. Sci. Res.*, vol. 11, no. 12, pp. 1424–1440, dic. 2022, doi: 10.21275/SR22123165037.

[30] M. Hasan and M. O. Faruq, "Ai-Augmented Risk Detection In Cybersecurity Compliance: A Grc-Based Evaluation In Healthcare And Financial," *ASRC Procedia Glob. Perspect. Sci. Scholarship*, vol. 01, no. 01, pp. 313–342, Jan. 2025, doi: 10.63125/49gs6175.

[31] A. N. John Wesly Sajja, "Enterprise Finance Reimagined: Harnessing ERP and Data Innovation for Next-Generation Value Creation," *Comput. Fraud Secur.*, vol. 2024, no. 4, p. 10, Apr. 2024, doi: 10.52710/cfs. 743.

[32] O. Sarioguz, E. Miser, and B. Teslim, "Redefining Governance, Risk, and Compliance (GRC) in the Digital Age: Integrating AI-Driven Risk Management Frameworks," *World J. Adv. Eng. Technol. Sci.*, vol. 10, no. 1, pp. 264–282, Oct. 2023, doi: 10.30574/wjaets. 2023.10.1.0257.

[33] O. Sarioguz, E. Miser, and B. Teslim, "Integrating AI in financial risk management: Evaluating the effects of machine learning algorithms on predictive accuracy and regulatory compliance," *Int. J. Sci. Res. Arch.*, vol. 13, no. 2, pp. 789–811, Nov. 2024, doi: 10.30574/ijsra.2024.13.2.2206.

[34] M. S. Soudi and M. Bauters, "AI Guidelines and Ethical Readiness Inside SMEs: A Review and Recommendations," *Digit. Soc.*, vol. 3, no. 1, p. 3, May 2024, doi: 10.1007/s44206-024-00087-1.

[35] S. H. Deep Kolagani, M. Bhandar, and R. Altounjy, "Integrating AI Predictive Analytics into Financial CRM for Retention," in *2025 7th International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, IEEE, Oct. 2025, pp. 1508–1513. doi: 10.1109/ICIDCA66325.2025.11280429.

[36] J. Smit, "Ethical AI in Banking: A Sectoral Literature Review with Case-Based Interventions," *Open J. Bus. Manag.*, vol. 13, no. 04, pp. 2420–2430, 2025, doi: 10.4236/ojbm.2025.134125.

[37] J. C. Ogeawuchi, A. Sharma, B. I. Adekunle, A. A. Abayomi, and O. Onifade, "Ethical Frameworks for AI Deployment in Financial Decision-Making: Balancing Profitability and Social Responsibility," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 10, no. 2, pp. 905–917, Apr. 2024, doi: 10.32628/CSEIT24102141.

[38] V. R. Gamit, "A Comparative Study of Ethical Governance of AI in Indian Public Sector Banks," *Int. J. Nov. Res. Dev.*, vol. 11, no. 1, pp. 489–496, 2026, doi: A Comparative Study of Ethical Governance of AI in Indian Public Sector Banks,".

[39] M. Madanchian and H. Taherdoost, "Ethical theories, governance models, and strategic frameworks for responsible AI adoption and organizational success," *Front. Artif. Intell.*, vol. 8, Jul. 2025, doi: 10.3389/frai.2025.1619029.

[40] A. Nastoska, B. Jancheska, M. Rizinski, and D. Trajanov, "Evaluating Trustworthiness in AI: Risks, Metrics, and Applications Across Industries," *Electronics*, vol. 14, no. 13, p. 2717, Jul. 2025, doi:

10.3390/electronics14132717.

[41] A. Vyas, “Revolutionizing Risk: The Role of Artificial Intelligence in Financial Risk Management, Forecasting, and Global Implementation,” 2025. doi: 10.2139/ssrn.5224657.

[42] H. Xie, Z. Tan, and X. Liu, “Application of Artificial Intelligence in Financial Risk Management in a Company,” in *Proceedings of the 2024 8th International Conference on Big Data and Internet of Things*, 2024, pp. 349–354. doi: 10.1145/3697355.3697412.

[43] M. El Hajj and J. Hammoud, “Unveiling the Influence of Artificial Intelligence and Machine Learning on Financial Markets: A Comprehensive Analysis of AI Applications in Trading, Risk Management, and Financial Operations,” *J. Risk Financ. Manag.*, vol. 16, no. 10, p. 434, Oct. 2023, doi: 10.3390/jrfm16100434.

[44] S Anamol, A Anjaneyulu, and M Raja Aseem, “Utilizing AI to predict and mitigate financial risks in banking and investment sectors,” *Int. J. Sci. Res. Arch.*, vol. 10, no. 1, pp. 1098–1104, oct. 2023, doi: 10.30574/ijsra.2023.10.1.0701.

[45] S. Fritz-Morgenthal, B. Hein, and J. Papenbrock, “Financial Risk Management and Explainable, Trustworthy, Responsible AI,” *Front. Artif. Intell.*, vol. 5, Feb. 2022, doi: 10.3389/frai.2022.779799.