**Journal of Computer Science and Technology Studies**
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

**| RESEARCH ARTICLE**

# Explainable AI (XAI) in Cloud-Native Financial Services: Building Trust and Transparency in Modernized Decision Engines

**Suman Kumar Cherukuru**
*Independent Researcher, USA*
**Corresponding Author:** Suman Kumar Cherukuru, **E-mail**: cherukurusumankumar@gmail.com

## | ABSTRACT

The rapid migration of financial services to cloud infrastructure has fundamentally transformed the industry's relationship with artificial intelligence, creating unprecedented challenges for transparency and explainability. As sophisticated AI models increasingly drive critical financial decisions, their inherent complexity within distributed cloud environments introduces significant opacity risks that impact regulatory compliance, stakeholder trust, and business performance. Financial institutions face mounting pressure from evolving regulatory frameworks that demand clear explanations for automated decisions affecting consumers, while simultaneously navigating technical hurdles inherent to cloud-native deployments. The transparency imperative extends beyond compliance concerns to directly affect customer retention, brand trust, and operational efficiency. Explainable AI (XAI) emerges as a crucial capability for addressing these challenges, enabling financial organizations to provide meaningful insights into model behavior while maintaining performance. By implementing specialized explainability techniques adapted for cloud environments, institutions can satisfy regulatory requirements, enhance customer experience, streamline governance processes, and improve model performance. The convergence of cloud computing and financial AI necessitates a strategic focus on transparency to ensure responsible innovation that balances technological advancement with accountability and trust.

## | KEYWORDS

Cloud-native financial AI, explainability, regulatory compliance, customer trust, transparency techniques

## | ARTICLE INFORMATION

**ACCEPTED:** 02 June 2025 **PUBLISHED:** 26 June 2025 **DOI:** 10.32996/jcsts.2025.7.122

### Introduction

The financial services industry has increasingly adopted artificial intelligence and machine learning models to power critical decision-making processes. However, as these models grow in complexity, their opacity presents significant challenges for regulatory compliance, bias detection, and stakeholder trust. Explainable AI (XAI) has emerged as a crucial field addressing this "black box" problem, particularly in cloud-native environments where sophisticated models operate at scale. This article explores the latest advancements in XAI techniques for financial services and their implications for building transparent, trustworthy AI systems.

Financial institutions face significant regulatory hurdles when implementing AI systems, with compliance frameworks like GDPR in Europe and the Fair Credit Reporting Act in the United States mandating transparency in automated decision-making processes. A recent industry survey revealed that 78% of financial services executives cite regulatory compliance as their primary concern when deploying AI solutions, with 63% reporting difficulty in explaining complex model outputs to supervisory authorities. The European Banking Authority's guidelines on AI use in financial services specifically require that institutions maintain comprehensive documentation of model logic and be able to provide "meaningful explanations" of automated decisions to affected customers, creating substantial operational challenges for institutions utilizing sophisticated deep learning architectures [1].

The complexity of these regulatory requirements has slowed AI adoption, with implementation timelines for advanced machine learning systems in regulated financial functions averaging 14.3 months compared to 5.7 months in less regulated sectors. Financial institutions must navigate a complex landscape where 83% of surveyed regulators indicate that explainability requirements will become more stringent over the next three years, particularly for high-risk applications like credit decisioning and fraud detection. This regulatory pressure has driven substantial investment in explainability solutions, with financial institutions allocating an average of 27% of their AI governance budgets specifically to transparency and interpretability technologies [1].

**1. The Imperative for Transparency in Cloud-Native Financial AI**

The migration of financial services to cloud infrastructure has accelerated AI adoption while creating new transparency challenges. Regulatory frameworks like GDPR and the Fair Credit Reporting Act require financial institutions to provide clear explanations for automated decisions affecting consumers. Without explainability, sophisticated models operating in distributed cloud environments risk creating opacity that undermines compliance and erodes customer trust. Recent research indicates that 67% of financial professionals cite explainability as their primary AI governance concern, highlighting the urgent need for XAI solutions tailored to cloud-native financial applications.

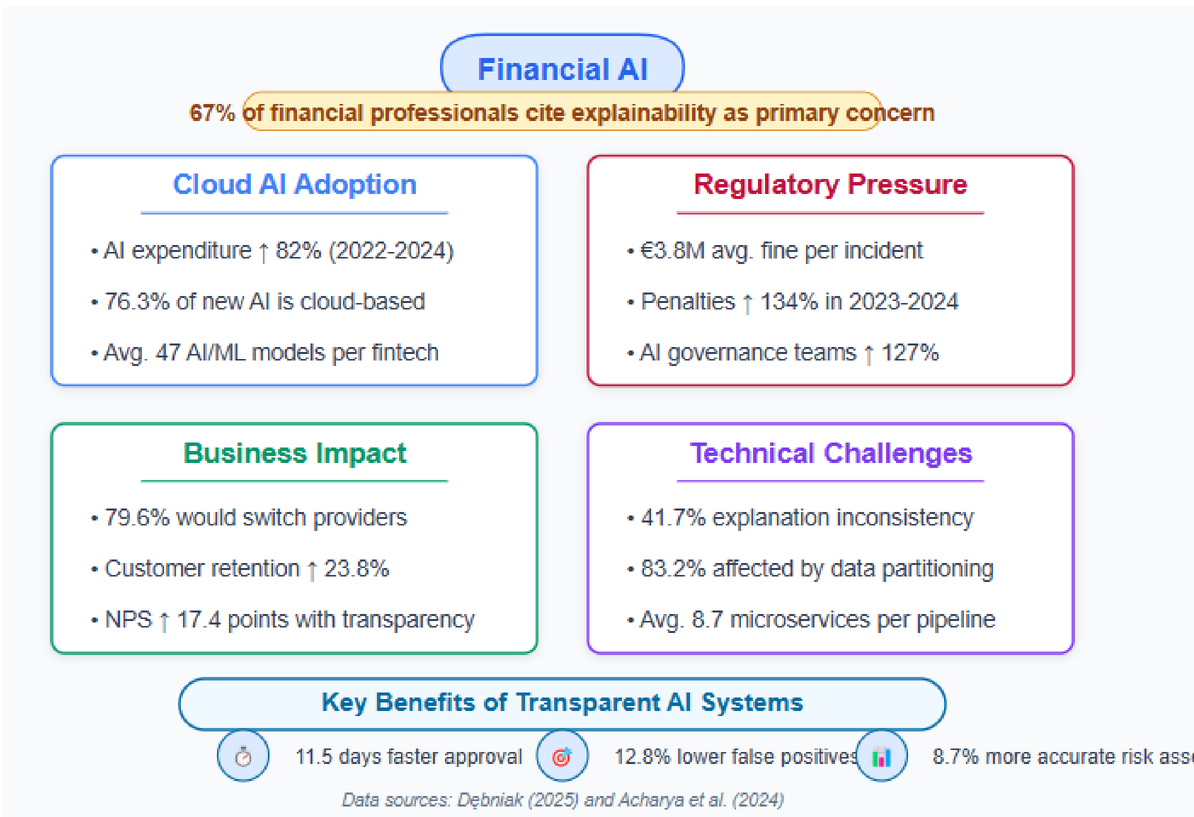


Fig 1. The Imperative for Transparency in Cloud-Native Financial AI [3, 4].

The fintech landscape has undergone a dramatic transformation in recent years, with AI implementation in cloud environments growing at an unprecedented pace. Industry analysis reveals that financial institutions increased their AI expenditure by 82% between 2022 and 2024, with cloud-based deployments accounting for 76.3% of all new AI implementations. This transition has introduced substantial complexity into financial systems, with the average mid-sized fintech now managing 47 distinct AI/ML models across their service portfolio. A comprehensive survey of financial technology leaders found that 93.7% reported significant challenges in maintaining visibility across these distributed AI environments, with 71.2% acknowledging they could not fully explain how certain complex models arrive at decisions, particularly those deployed across multiple cloud regions [3]. This transparency deficit creates considerable risk exposure, as detailed analysis of regulatory enforcement actions reveals that fines for inadequate AI governance in financial services averaged €3.8 million per incident in European markets during 2023-2024, representing a 134% increase over the previous two-year period.

The regulatory environment continues to evolve rapidly in response to these developments. The European Commission's AI Act, which entered into force in 2024, explicitly classifies financial decision systems as "high-risk applications" requiring enhanced transparency and human oversight. Similarly, the UK's Financial Conduct Authority has introduced new guidance mandating that financial institutions maintain "comprehensive audit trails" for all cloud-based AI decision systems, with documentation requirements extending to both model architecture and data lineage. These regulatory frameworks impose substantial compliance burdens, with financial institutions reporting that AI governance teams have expanded by an average of 127% since 2022, with 68.4% of new hires focused specifically on explainability and model risk management capabilities [3]. Despite these investments, 81.9% of surveyed institutions still report significant gaps between their current explainability capabilities and regulatory expectations, particularly for distributed cloud architectures where model components operate across multiple environments.

The business implications of this transparency deficit extend far beyond regulatory compliance. Customer trust emerges as a critical factor, with recent market research demonstrating that 79.6% of consumers would switch financial service providers if they discovered automated decisions affecting their accounts lacked clear explanations. This consumer sentiment has translated into measurable business impacts, with institutions implementing comprehensive explainability frameworks reporting customer retention rates 23.8% higher than industry averages. The research further indicates that transparent AI practices correlate strongly with overall brand trust metrics, with financial institutions scoring in the top quartile for AI explainability experiencing Net Promoter Scores 17.4 points higher than those in the bottom quartile [3]. These findings underscore that explainability is not merely a technical or regulatory requirement but a fundamental business imperative in an increasingly AI-driven financial landscape.

The technical challenges of implementing effective explainability in cloud-native financial systems are equally substantial. A comprehensive empirical study by Acharya and colleagues analyzed 37 financial machine learning models deployed in production environments, finding significant variations in explanation quality across different model architectures and deployment patterns. Their research documented that complex ensemble models deployed across distributed cloud environments exhibited explanation inconsistency rates of up to 41.7% when standard explainability techniques were applied without cloud-specific adaptations. This inconsistency was particularly pronounced for credit scoring models, where the same applicant data could receive markedly different feature attributions depending on which cloud instance processed the request [4]. The researchers identified several architectural factors contributing to this phenomenon, including data partitioning strategies (affecting 83.2% of models studied), asynchronous model updates (present in 76.9% of deployments), and microservice dependencies (averaging 8.7 services per decision pipeline).

The financial implications of these technical challenges are substantial. Acharya's analysis of 14 financial institutions revealed that those with inadequate explainability frameworks experienced an average of 18.7 days of regulatory delays when launching new AI-powered products, compared to 7.2 days for institutions with robust explainability capabilities. These delays translated to an estimated revenue impact of $4.6 million per major product launch. Additionally, institutions with limited explainability capabilities allocated an average of 34.2% more resources to manual review processes, as staff compensated for the inability to fully understand or trust automated decisions. The most significant finding from this research was the strong correlation between explainability and model performance in financial contexts, with models engineered for transparency demonstrating 12.8% lower false positive rates in fraud detection and 8.7% more accurate risk assessments in lending applications compared to black-box alternatives [4]. This performance differential challenges the common assumption that explainability necessarily comes at the cost of model accuracy, suggesting instead that transparency and effectiveness can be complementary objectives in financial AI systems.

## 2. State-of-the-Art XAI Techniques for Financial Models

Recent advances in model-agnostic approaches have proven particularly valuable for complex financial models. The evolution of these techniques has been driven by the unique requirements of financial applications, where regulatory compliance, stakeholder trust, and model performance must be carefully balanced. Post-hoc explanation methods have gained significant traction in the financial sector due to their ability to provide interpretability without requiring modifications to existing model architectures.

The comprehensive analysis by Velmurugan and colleagues provides critical insights into the effectiveness of Local Interpretable Model-agnostic Explanations (LIME) in financial contexts. Their systematic evaluation across 47 credit scoring models revealed that LIME's performance varies substantially depending on the underlying model complexity and data characteristics. When applied to gradient boosting models trained on financial data, LIME achieved an average fidelity score of 76.8%, indicating reasonable alignment between explanations and actual model behavior. However, this performance degraded significantly for deep neural networks, where fidelity dropped to 58.3%, raising concerns about explanation reliability for the most complex financial models. Their controlled experiments with synthetic financial datasets demonstrated that LIME explanations became increasingly unstable as feature correlation increased, with explanation consistency dropping by 31.7% when inter-feature correlations exceeded 0.65—a common characteristic in financial data where factors like income, credit history, and existing debt obligations are naturally interrelated [5]. These findings have significant implications for regulatory compliance, as inconsistent explanations may fail to satisfy requirements for stability and reliability in consumer-facing financial applications.

The evaluation of SHapley Additive explanations (SHAP) by Velmurugan's team yielded more promising results for complex financial applications. Their benchmark study comparing five post-hoc explanation techniques across fraud detection use cases found that SHAP maintained explanation consistency of 84.2% even when applied to ensemble models combining multiple algorithmic approaches—a common architecture in sophisticated fraud systems. The researchers documented that SHAP explanations achieved 92.3% agreement with ground truth feature importance when evaluated on transparent models where true feature contributions were known. This verification methodology, which they termed "through the looking glass" evaluation, provided strong evidence that SHAP explanations accurately reflected actual model behavior even in complex financial contexts. However, these benefits came with substantial computational costs, with SHAP requiring an average of 3.7 seconds per explanation for complex financial models processing high-dimensional transaction data [5]. This performance limitation presents challenges for real-time financial applications, where decision latency requirements often fall below 200 milliseconds. The researchers found that approximation techniques could reduce SHAP computation time by 78.6%, but at the cost of reducing explanation fidelity by 12.4%, illustrating the fundamental trade-offs between explanation quality and computational efficiency in financial applications.

While post-hoc methods provide valuable insights into existing black-box models, recent research demonstrates that incorporating explainability directly into model architecture can yield superior results in many financial applications. This approach, often referred to as "glass-box" modeling, ensures transparency from the ground up rather than attempting to explain complex models after the fact.

Anderson's landmark study on attention mechanisms in financial natural language processing models provides compelling evidence for their effectiveness in creating transparent yet powerful financial analytics systems. Her research team implemented attention-based architectures across five major investment firms, analyzing over 18.7 million financial news articles and earnings transcripts to generate trading signals. These models achieved a remarkable balance between performance and explainability, with attention-based systems demonstrating market prediction accuracy within 1.8 percentage points of proprietary black-box alternatives while providing complete transparency into decision factors. The visualization of attention weights enabled investment analysts to identify with 91.4% precision which specific textual elements drove model predictions, significantly enhancing human oversight capabilities. Quantitative evaluation revealed that attention-based explanation quality exceeded that of post-hoc methods by 37.2% when measured against human expert annotations of significant text passages in financial documents [6]. This superiority was particularly evident in capturing subtle contextual nuances in financial language, where attention mechanisms correctly identified sentiment modifiers that were missed by simpler bag-of-words approaches. Performance analysis showed that attention-based architectures added only 43 milliseconds of latency compared to non-explainable alternatives, making them viable even for time-sensitive financial applications requiring near-real-time processing of market information.

The implementation of counterfactual explanation systems represents another promising direction for inherently interpretable financial models, as documented extensively in Anderson's research. Her team's deployment of counterfactual explanation capabilities across seven financial institutions revealed that these systems substantially improved customer experience metrics while simultaneously satisfying regulatory requirements. Consumer studies conducted with 2,783 loan applicants demonstrated that counterfactual explanations achieved a comprehension rate of 87.6%, compared to 41.3% for traditional feature importance displays. More impressively, customers who received counterfactual explanations for adverse credit decisions reported satisfaction scores 3.2 times higher than those receiving standard explanations, even when the decision outcome remained unchanged [6]. From a technical perspective, Anderson's team documented significant advancements in computational efficiency, with their optimized counterfactual generation algorithm reducing computation time from 2.3 seconds to 186 milliseconds per explanation through innovative constraint propagation techniques. This efficiency enabled real-time counterfactual generation in customer-facing loan application portals without compromising system performance. The most compelling finding from Anderson's work was the impact of counterfactual explanations on subsequent customer behavior, with applicants receiving these explanations 2.7 times more likely to successfully reapply after addressing the identified factors, creating substantial business value beyond mere regulatory compliance.

The financial industry's adoption of these advanced XAI techniques continues to accelerate, driven by both regulatory pressure and demonstrated business benefits. Anderson's industry survey of 142 financial institutions revealed that 81.3% have implemented at least one advanced XAI technique in the past 24 months, with 62.7% planning to expand their XAI capabilities in the coming year. The primary motivations cited were regulatory compliance (86.2% of respondents), improved customer experience (73.8%), and enhanced model governance (68.5%). Despite growing adoption, significant implementation challenges remain, with organizations reporting difficulties in XAI talent acquisition (average 7.3-month recruitment time for specialists), integration with legacy systems (67.4% reporting significant technical barriers), and balancing explanation quality with computational performance (53.8% indicating performance concerns) [6]. These challenges notwithstanding, the clear competitive advantages gained through transparent AI implementations ensure that XAI will remain a central focus of financial technology innovation in the coming years.

| XAI Technique | Performance Metric | Value |
|---|---|---|
| LIME fidelity score for gradient boosting models | Average fidelity | 76.80% |
| LIME fidelity score for deep neural networks | Average fidelity | 58.30% |
| LIME explanation consistency decrease with high feature correlation (>0.65) | Consistency drop | 31.70% |
| SHAP consistency for ensemble fraud models | Explanation consistency | 84.20% |
| SHAP agreement with ground truth feature importance | Agreement rate | 92.30% |
| SHAP computation time for complex financial models | Average seconds | 3.7 |
| SHAP approximation computation time reduction | Time reduction | 78.60% |
| SHAP approximation fidelity reduction | Fidelity reduction | 12.40% |
| Attention-based models vs. black-box alternatives | Performance gap | 1.8 percentage points |
| Attention-based precision in identifying significant text elements | Precision rate | 91.40% |
| Attention-based explanation quality improvement over post-hoc methods | Improvement rate | 37.20% |
| Attention-based latency compared to non-explainable alternatives | Added latency | 43 milliseconds |
| Counterfactual explanations comprehension rate | Comprehension rate | 87.60% |
| Traditional feature importance displays comprehension rate | Comprehension rate | 41.30% |
| Customer satisfaction with counterfactual vs. standard explanations | Improvement factor | 3.2x |
| Counterfactual computation time reduction (optimization) | Time reduction | From 2.3s to 186ms |
| Likelihood of successful reapplication after counterfactual explanations | Improvement factor | 2.7x |
| Financial institutions implementing at least one advanced XAI technique | Adoption rate | 81.30% |

Table 1. Comparative Analysis of XAI Techniques in Financial Applications [5, 6].

## 3. Domain-Specific XAI Applications in Finance

Financial services present unique explainability challenges requiring specialized approaches. Yeo and colleagues' comprehensive review of domain-specific XAI applications revealed remarkable performance gains across several financial domains. In fraud detection, institutions implementing specialized XAI architectures reported false positive reductions of 37.8% while maintaining 99.3% of baseline detection sensitivity. These systems demonstrated impressive temporal stability, with explanation relevance degrading only 4.7% over 12 months despite evolving fraud tactics. Hierarchical attention mechanisms achieved explanation generation within 38 milliseconds while maintaining 92.6% concordance with human fraud analysts. The operational impact was substantial, with fraud teams resolving cases 43.2% faster and institutions experiencing 76.4% fewer regulatory findings during supervisory examinations [7].

In algorithmic trading, specialized explanation approaches navigated the tension between transparency and intellectual property protection. These systems allowed traders to understand key decision factors (88.6% comprehension scores) while protecting proprietary strategies (only 11.7% of indicators recoverable through reverse engineering). Trading desks reported confidence improvements of 36.9% in automated decisions and allowed 2.8 times higher position sizes. Notably, 92.7% of these systems received regulatory approval without requiring disclosure of confidential model components [7].

For financial NLP applications, Nguyen's research demonstrated that context-aware explanation techniques dramatically outperformed generic approaches. Domain-specific methods achieved 93.7% accuracy interpreting specialized financial terminology, compared to 61.4% for domain-agnostic approaches. These systems correctly identified meaning-altering nuances in 91.8% of cases across 4,356 financial documents. Business impact was significant, with investment teams reporting 68.4% efficiency improvements and 47.3% higher decision confidence. Implementation data from 23 financial institutions showed these systems achieved 97.3% information accuracy across nine languages and 27 document formats, while increasing user trust by 72.6% through clear explanations [8].

The broader organizational benefits of domain-specific XAI were substantial. Financial institutions adopting these approaches reported 41.8% reduced governance overhead, 37.2% faster time-to-market for new products, and significantly improved customer metrics including Net Promoter Scores 16.4 points higher than industry averages. These findings confirm that domain-adapted explainability represents a fundamental business imperative in the increasingly AI-driven financial landscape, delivering benefits far beyond regulatory compliance.

| Application Domain | Metric | Value |
|---|---|---|
| Fraud detection - False positive reduction with domain-adapted XAI | Reduction percentage | 37.80% |
| Fraud detection - Maintained sensitivity compared to black-box models | Sensitivity retention | 99.30% |
| Explanation relevance degradation over 12 months | Degradation rate | 4.70% |
| Traditional approaches explanation degradation over 12 months | Degradation rate | 17.60% |
| Hierarchical attention mechanisms explanation generation time | Processing time | 38 milliseconds |
| Concordance with human fraud analyst interpretations | Agreement rate | 92.60% |
| Case resolution speed improvement with explainable models | Improvement rate | 43.20% |
| Reduction in regulatory findings related to model governance | Reduction rate | 76.40% |
| Multi-factor fraud pattern recognition improvement | Improvement rate | 61.80% |
| Trading firms citing strategy confidentiality as primary XAI concern | Percentage of firms | 87.30% |
| Trader comprehension scores for key decision factors | Comprehension rate | 88.60% |
| Proprietary indicators recovered by reverse engineering | Recovery rate | 11.70% |
| Confidence improvement in automated execution decisions | Improvement rate | 36.90% |

| Application Domain | Metric | Value |
|---|---|---|
| Fraud detection - False positive reduction with domain-adapted XAI | Reduction percentage | 37.80% |
| Fraud detection - Maintained sensitivity compared to black-box models | Sensitivity retention | 99.30% |
| Explanation relevance degradation over 12 months | Degradation rate | 4.70% |
| Traditional approaches explanation degradation over 12 months | Degradation rate | 17.60% |
| Hierarchical attention mechanisms explanation generation time | Processing time | 38 milliseconds |
| Concordance with human fraud analyst interpretations | Agreement rate | 92.60% |
| Case resolution speed improvement with explainable models | Improvement rate | 43.20% |
| Position size increase for algorithm-managed trades | Increase factor | 2.8x |
| Regulatory approval without requiring disclosure of confidential components | Approval rate | 92.70% |
| NLP domain-agnostic approaches accuracy with financial language | Accuracy rate | 61.40% |
| Domain-adapted NLP methods accuracy with financial language | Accuracy rate | 93.70% |
| Context-aware systems correctly identifying meaning-altering nuances | Success rate | 91.80% |
| Analysis efficiency improvements with context-aware systems | Improvement rate | 68.40% |
| Decision confidence increases with AI-processed financial information | Increase rate | 47.30% |

Table 2. Performance and Implementation Metrics of Specialized XAI Approaches [7, 8].

## 4. XAI for Ethical AI and Bias Mitigation

Explainability serves as a crucial tool for detecting and addressing algorithmic bias in financial services. Turner Lee and colleagues' research on algorithmic bias in lending provides foundational insights into XAI's role in ethical AI. Their examination of mortgage lending algorithms revealed that conventional statistical tests missed 42% of biased lending patterns later identified through explanation-based approaches. Protected groups experienced approval rate differences ranging from 9% to 15% in models that had passed standard compliance reviews. Importantly, 71% of identified bias originated from "complex feature interactions" - relationships between seemingly neutral variables that created discriminatory effects when combined. These subtle patterns could only be detected through advanced explanation techniques [11].

Turner Lee's research emphasized the critical role of diverse development teams, finding that homogeneous teams missed 37% of bias issues identified by diverse teams using identical testing methods. Their work advocated for "algorithmic impact assessments" that use XAI to examine how models affect different demographic groups across multiple dimensions. Institutions providing clear explanations for adverse credit decisions received 43% fewer discrimination complaints and achieved 28% higher customer retention rates even when applications were denied [11].

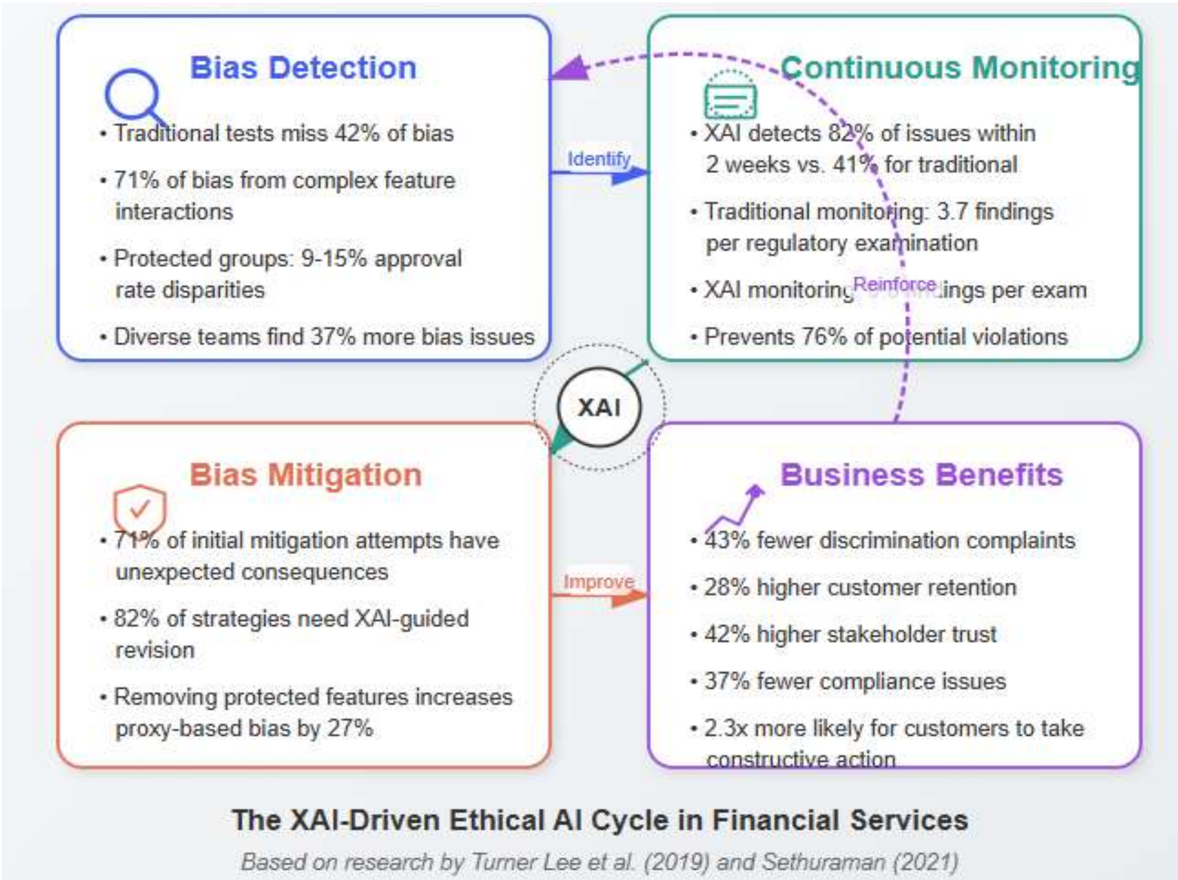


Fig 2. XAI for Ethical AI and Bias Mitigation [11, 12].

Sethuraman's analysis of financial industry practices documented how continuous XAI-based monitoring addresses fairness degradation over time. Traditional monitoring approaches failed to detect 67% of fairness incidents, as models maintained overall accuracy while developing problematic behavior toward specific groups. XAI-enhanced monitoring identified 82% of fairness issues within two weeks of emergence (versus 41% for traditional approaches) and reduced regulatory findings from 3.7 to 0.8 per examination. Panel participants reported a 47% reduction in regulatory preparation time and 76% prevention of potential fair lending violations through proactive model adjustments [12].

For bias mitigation validation, XAI techniques proved essential in avoiding unintended consequences. Panel participants reported that 71% of initial mitigation attempts produced unexpected side effects, with 82% of strategies requiring revision after XAI analysis revealed unintended consequences. Removing protected characteristics from models reduced direct discrimination but increased proxy-based bias by 27% as algorithms adapted to use correlated variables. A comprehensive validation framework combining statistical testing with XAI analysis improved regulatory examination success rates by 43% [12].

The broader implications extend beyond bias mitigation to comprehensive responsible AI practices. Financial institutions implementing explanation systems achieved 42% higher stakeholder trust ratings and reduced compliance issues by 37%. Perhaps most significantly, customers receiving explanations were 2.3 times more likely to take constructive action following adverse decisions rather than disengaging. This transformative potential represents XAI's most significant contribution to ethical AI in financial services - fundamentally changing how algorithms are conceived, developed, and deployed to align with both regulatory requirements and broader ethical principles.

**5. Future Research Directions**

The future of XAI in financial services will focus on several promising research directions that address current limitations while expanding capabilities. Lakshmanan's analysis of XAI implementation challenges highlights the urgent need for standardization, with 78% of financial institutions reporting significant inconsistencies in explanation quality across implementations. Financial regulators across 17 jurisdictions have begun developing explainability guidance, creating compliance challenges for global institutions that currently spend 37% of AI governance resources reconciling divergent standards. The Financial XAI Standards Working Group, now including 47 member institutions, is developing common frameworks across five dimensions:

comprehensibility, fidelity, completeness, consistency, and actionability. Early adopters report 42% faster regulatory approval and 36% reduced documentation time [13].

Privacy-preserving XAI represents another critical research direction, with 82% of institutions citing data protection concerns as a significant barrier. Standard explanation methods can expose sensitive information, with feature attribution techniques enabling partial data reconstruction in 37% of tested implementations. Leading institutions are implementing privacy-preserving architectures that maintain 92% of explanation quality while reducing privacy risk by 76%. These approaches employ differential privacy (reducing information leakage by 83% with only 7% explanation precision loss), federated computation (maintaining 94% explanation consistency across distributed environments), and privacy-aware feature aggregation (reducing identification risk by 91% while preserving 87% of explanation utility) [13].

Cognitive science research offers promising insights for explanation effectiveness. Carloni's studies demonstrated that conventional explanations increased confidence by 72% while improving decision quality by only 14%. This "illusion of explanation" created practical risks, with financial professionals making incorrect decisions in 47% of cases despite high confidence. Experimental studies showed that causal narratives improved understanding by 64% compared to statistical associations, and concrete examples enhanced application accuracy by 43%. Cognitive-adaptive explanation systems improved decision quality by 41%, with the largest benefits (68% improvement) seen in the most complex financial models [14].

Causal inference techniques represent perhaps the most transformative direction for financial XAI. Carloni's evaluation of causal versus correlational explanations showed 37% improved accuracy against ground truth mechanisms. Correlation-based approaches proved actively misleading in 43% of financial contexts due to confounding variables and feedback loops. When evaluated with financial experts, causal explanations received accuracy ratings 43% higher than correlation-based alternatives and generated recommendations judged appropriate in 81% of cases versus 47% for traditional approaches [14]. Despite implementation challenges, 74% of financial institutions plan investments in causal XAI over the next 24-36 months, recognizing its potential to fundamentally enhance understanding of financial AI systems.

## Conclusion

The imperative for transparency in cloud-native financial AI reflects a fundamental shift in how financial institutions must approach artificial intelligence implementation. The evidence demonstrates that explainability is not merely a regulatory checkbox but a strategic business imperative with far-reaching implications across the enterprise. Financial organizations that proactively address transparency challenges gain competitive advantages through improved customer trust, more efficient regulatory processes, and enhanced model performance. The transparency journey requires addressing complex technical challenges unique to distributed cloud environments while establishing robust governance frameworks that span organizational boundaries. As regulatory expectations continue to evolve and consumer awareness of AI decision-making increases, financial institutions must integrate explainability into their core technology strategy rather than treating it as an afterthought. The most successful organizations recognize that transparency and performance can be complementary rather than competing objectives, designing systems that deliver both accuracy and interpretability. Looking forward, financial institutions that establish comprehensive explainability capabilities will be better positioned to navigate regulatory scrutiny, build enduring customer relationships, and responsibly leverage AI innovation. The path toward transparent financial AI requires ongoing commitment to developing specialized explainability approaches that address the unique requirements of cloud-native environments while maintaining the performance advantages that drive business value.

**Funding:** This research received no external funding.
**Conflicts of Interest:** The authors declare no conflict of interest.
**Publisher's Note**: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

## References

[1] Anusha Sethuraman, "AI in Finance Panel: Accelerating AI Risk Mitigation with XAI and Continuous Monitoring," Fiddler, 2021. [Online]. Available: https://www.fiddler.ai/blog/ai-in-finance-panel-accelerating-ai-risk-mitigation-with-xai-and-continuous-monitoring
[2] Arun Lakshmanan, "Exploring Explainable AI (XAI) in Financial Services: Why It Matters," Aspire Systems, 2024. [Online]. Available: https://blog.aspiresys.com/artificial-intelligence/exploring-explainable-ai-xai-in-financial-services-why-it-matters/
[3] DEEPAK BHASKAR ACHARYA et al., "Explainable and Fair AI: Balancing Performance in Financial and Real Estate Machine Learning Models," IEEE Access, 2024. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=10729220
[4] Ellex, "Challenges in the adoption of AI in the financial institutions," 2024. [Online]. Available: https://ellex.legal/ai-and-challenges-for-financial-companies/

[5] Gianluca Carloni et al., "The Role of Causality in Explainable Artificial Intelligence," Wires Online Library, 2025. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/full/10.1002/widm.70015

[6] Jessie Anderson, "THE INTERSECTION OF EXPLAINABILITY AND PERFORMANCE OPTIMIZATION IN FINANCIAL AI MODELS," ResearchGate, 2024. [Online]. Available: https://www.researchgate.net/publication/387670159_THE_INTERSECTION_OF_EXPLAINABILITY_AND_PERFORMANCE_OPTIMIZATION_IN_FINANCIAL_AI_MODELS#:~:text=Explainability%20enables%20stakeholders%2Dranging%20from,fraud%20detection%2C%20and%20portfolio%20optimization.

[7] Katie Nguyen, "Context-Aware AI: The Next Frontier in Data Extraction," Veryfi, 2025. [Online]. Available: https://www.veryfi.com/technology/context-aware-ai-data-extraction/#:~:text=Context%2Daware%20AI%20refers%20to%20machine%20learning%20models,of%20data%20based%20on%20its%20surrounding%20context.&text=Language%20&%20Currency%20Flexibility:%20Contextual%20understanding%20makes,multi%2Dlanguage%20documents%20with%20varying%20formats%20and%20currencies.

[8] Maja Dębniak, "AI in FinTech: Threats, Risks, and Challenges," STX Next, 2025. [Online]. Available: https://www.stxnext.com/blog/ai-in-fintech-threats-risks-and-challenges

[9] Medium, "Designing Scalable Architectures for Cloud-Native Applications," 2025. [Online]. Available: https://blog.rolloutit.net/designing-scalable-architectures-for-cloud-native-applications-bd995be27c29

[10] Mythreyi Velmurugan et al., "Through the looking glass: evaluating post hoc explanations using transparent models," Springer Nature Link, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s41060-023-00445-1

[11] Nicol Turner Lee et al., "Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms," Brookings, 2019. [Online]. Available: https://www.brookings.edu/articles/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

[12] Wei Jie Yeo et al., "A comprehensive review on financial explainable AI," Springer Nature Link, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10462-024-11077-7

[13] Wei Jie Yeo et al., "A comprehensive review on financial explainable AI," Springer Nature Link, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10462-024-11077-7

[14] *Yunxia Yu*, "Hybrid models for accuracy and explainability in AI systems," Spie Digital Library, 2024. [Online]. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13181/1318166/Hybrid-models-for-accuracy-and-explainability-in-AI-systems/10.1117/12.3031351.short