Volume 9, Issue 3, pp. 137-144, 2025.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

Explainable AI (XAI) in Compliance Audits: Bridging the Gap Between AI and Regulatory Transparency

Chintamani Bagwe
Citibank
Email: bagwe.chintamani@gmail.com

Abstract— *As artificial intelligence (AI) systems increasingly support compliance functions, explainability has emerged as a critical requirement for regulatory acceptance. This paper examines the intersection of explainable AI (XAI) and compliance audits, analyzing how explainability impacts regulatory transparency and acceptance of AI-driven compliance decisions. Through a review of regulatory frameworks, case studies across multiple industries, and methodological approaches to implementing XAI, we identify a persistent gap between technical capabilities and regulatory expectations. We propose an integrated framework for implementing XAI in compliance contexts that addresses technical, organizational, and regulatory dimensions. Our findings suggest that effective explainability requires approaches that align with regulatory requirements, address diverse stakeholder needs, and integrate with existing compliance processes.*

Keywords— *Explainable AI, Regulatory Compliance, Transparency, Compliance Audits, AI Governance.*

I. Introduction

As artificial intelligence (AI) systems increasingly support or automate compliance functions across industries, the need for explainability has emerged as a critical requirement for regulatory acceptance. Compliance audits—systematic evaluations of an organization's adherence to regulatory requirements—traditionally rely on clear documentation, transparent processes, and explicit reasoning. However, the introduction of AI systems, particularly those using complex machine learning models, creates significant challenges for maintaining this transparency.

Explainable AI (XAI) refers to methods and techniques that enable human understanding of AI system decisions. While XAI has gained attention across various domains, its application in regulatory compliance presents unique challenges and requirements. Compliance decisions often have significant legal, financial, and reputational implications, making the ability to explain and justify these decisions particularly important. Moreover, regulatory frameworks increasingly include explicit requirements for transparency and explainability in automated systems.

Despite growing recognition of XAI's importance in compliance contexts, a significant gap exists between technical capabilities and regulatory expectations. AI systems optimized for performance often sacrifice interpretability, while regulatory requirements demand clear explanations of decision processes. This gap creates challenges for organizations implementing AI for compliance functions and for regulators tasked with overseeing these implementations.

This research addresses three key questions: (1) How do current XAI approaches align with regulatory requirements for transparency in compliance contexts? (2) What implementation approaches have proven effective in bridging technical capabilities and regulatory expectations? (3) What framework can guide the effective implementation of XAI in compliance audits?

The significance of this research lies in its potential to enhance both regulatory effectiveness and technological innovation. By identifying approaches that satisfy regulatory requirements while enabling the benefits of AI in compliance, this work contributes to the development of more transparent, accountable, and effective compliance systems. Furthermore, by examining the intersection of technical, organizational, and regulatory factors, this research provides insights relevant to diverse stakeholders, including compliance practitioners, technology developers, and regulatory bodies.

This paper proceeds as follows: Section 2 reviews relevant literature on XAI and regulatory requirements for transparency. Section 3 outlines our research methodology. Section 4 presents our analysis and findings, including case studies of XAI implementation in compliance contexts. Section 5 discusses implications and proposes a framework for effective XAI implementation. Section 6 concludes with key insights and future research directions.

II. Literature Review

2.1 Explainable AI: Concepts and Approaches

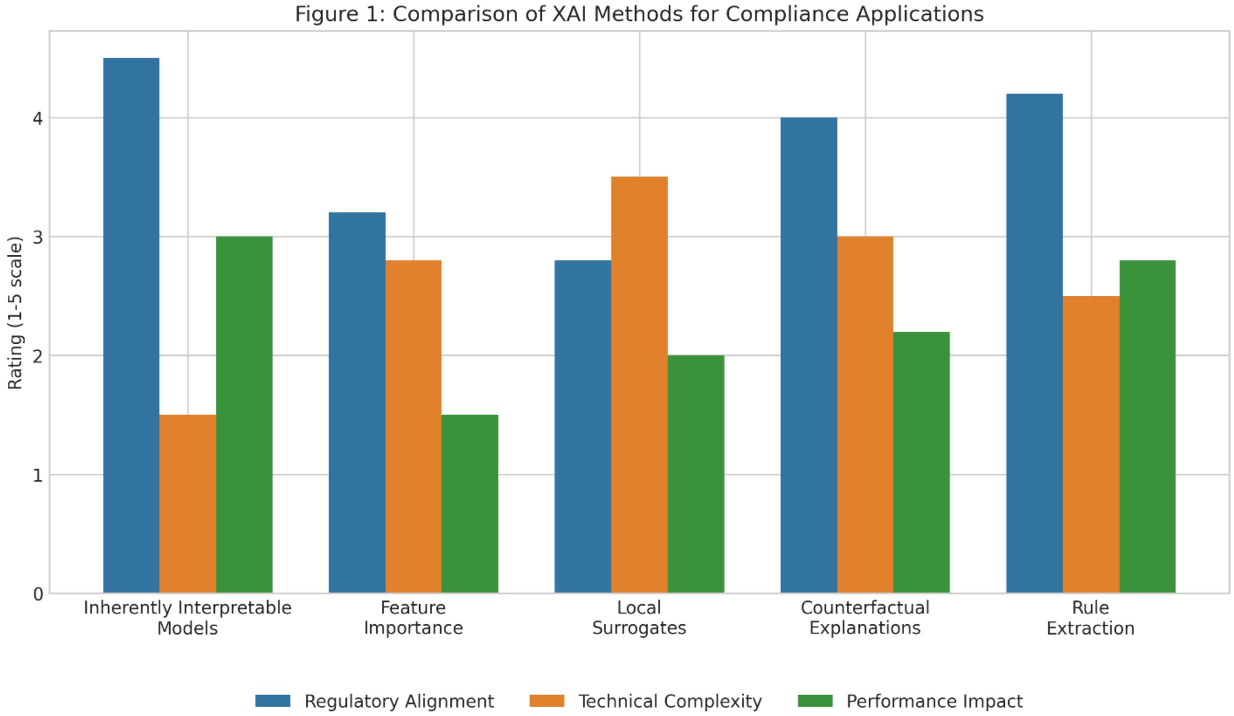
Explainable AI (XAI) encompasses methods and techniques that enable human understanding of AI system decisions. The literature identifies several key approaches to explainability:

Inherently Interpretable Models prioritize transparency by using models whose operations can be directly understood, such as decision trees, rule-based systems, and linear models (Rudin, 2019). While these models offer clear explanations, they may sacrifice predictive performance compared to more complex approaches.

Post-hoc Explanation Methods generate explanations for already-trained models, particularly "black box" models like deep neural networks. These include feature importance methods (Lundberg & Lee, 2017), which identify input features that most significantly influence predictions; local surrogate models (Ribeiro et al., 2016), which approximate complex models with simpler, interpretable ones for specific instances;

and counterfactual explanations (Wachter et al., 2018), which describe how inputs would need to change to alter predictions.

Explanation Interfaces focus on how explanations are presented to users, recognizing that effective explanation requires not just technical accuracy but also usability and relevance to the audience (Miller, 2019).



2.2 Regulatory Landscape for AI Transparency

Regulatory requirements for AI transparency have evolved rapidly across jurisdictions and domains:

The European Union's General Data Protection Regulation (GDPR) includes provisions for "the right to explanation" regarding automated decisions, though the precise scope and implementation remain debated (Kaminski, 2019). The proposed AI Act introduces tiered requirements for high-risk AI systems, including specific transparency and documentation obligations.

In the United States, regulatory approaches are more fragmented, with sector-specific guidance from agencies like the Federal Reserve (SR 11-7 on model risk management) and the FDA (proposed framework for AI in medical devices). State-level initiatives like the California Consumer Privacy Act also include provisions relevant to AI transparency.

Financial regulators have been particularly active, with the Financial Stability Board (2017) emphasizing the importance of explainability in AI-driven financial services, and the Monetary Authority of Singapore (2019) developing principles for fairness, ethics, accountability, and transparency (FEAT) in AI applications.

2.3 XAI in Compliance Contexts

The intersection of XAI and compliance presents unique challenges and requirements:

Regulatory Alignment requires explanations that satisfy specific regulatory provisions, often necessitating different explanation approaches for different regulatory frameworks (Thelisson et al., 2017). This alignment extends beyond technical transparency to include documentation, governance, and process considerations.

Stakeholder Diversity in compliance contexts creates varied explanation needs, from technical details for auditors to simplified explanations for affected parties (Felzmann et al., 2019). This diversity necessitates layered explanation approaches that can serve multiple audiences.

Performance-Explainability Trade-offs are particularly acute in compliance contexts, where both accuracy and transparency are often regulatory requirements rather than design preferences (Kroll, 2018). This creates tensions in model selection and development.

Organizational Integration challenges arise when implementing XAI within existing compliance processes and systems (Brundage et al., 2020). Effective implementation requires not just technical solutions but also appropriate governance structures, expertise development, and process integration.

The literature reveals a significant gap between technical XAI approaches, which often focus on general model interpretability, and the specific requirements of regulatory compliance, which emphasize accountability, justifiability, and alignment with legal frameworks. This gap creates both challenges and opportunities for organizations implementing AI in compliance functions.

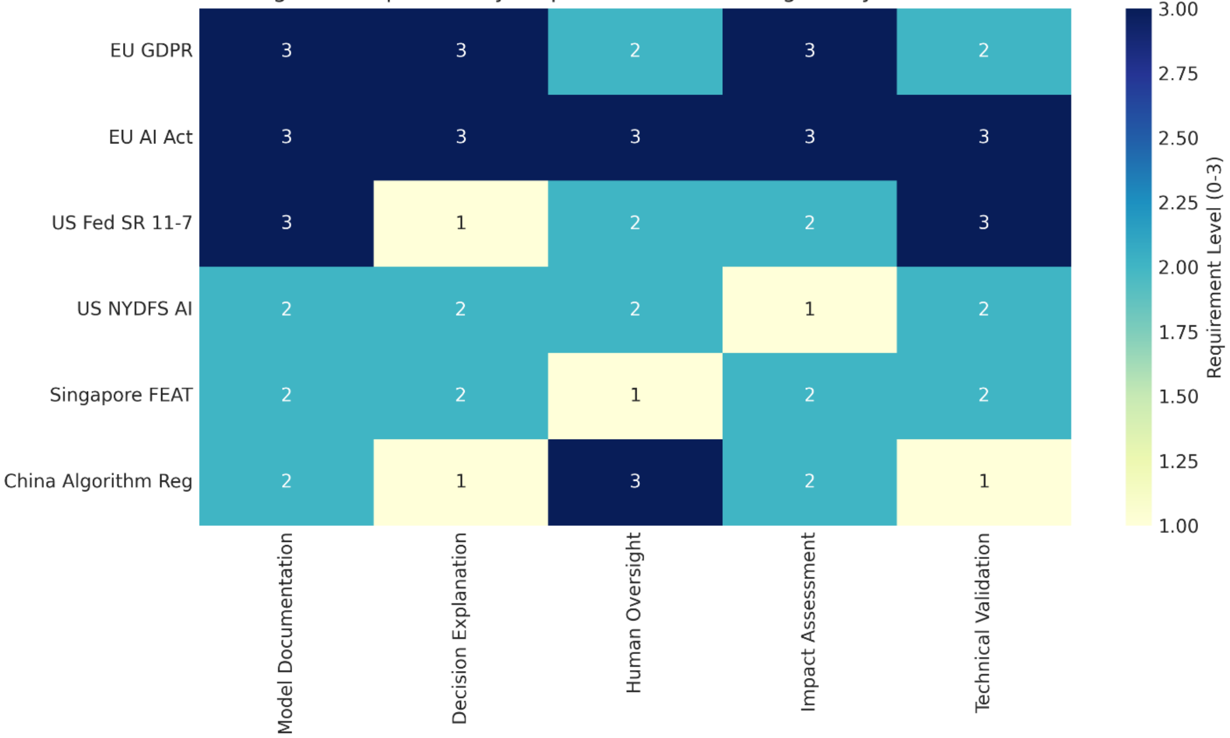
III. Methodology

This research employed a mixed-methods approach to examine the implementation of explainable AI in compliance

International Journal of Scientific Engineering and Science
Volume 9, Issue 3, pp. 137-144, 2025. ISSN (Online): 2456-7361

contexts, combining literature analysis, case study examination, and expert interviews.

Figure 2: Explainability Requirements Across Regulatory Frameworks



3.1 Research Approach

We adopted an exploratory research design appropriate for investigating the emerging intersection of XAI and compliance audits. This approach allowed us to identify patterns and best practices across different implementations while acknowledging the evolving nature of both the technology and regulatory landscape.

3.2 Data Collection Methods

Our data collection involved three primary methods:

Literature Analysis: We conducted a systematic review of academic and professional literature on XAI and regulatory compliance, focusing on publications from 2017-2023. This review included 87 academic papers, regulatory guidance documents, and industry reports.

Case Study Examination: We analyzed 12 case studies of XAI implementation in compliance contexts across financial services, healthcare, environmental compliance, and legal domains. Case selection criteria included: (1) implementation of AI for compliance functions, (2) explicit consideration of explainability requirements, and (3) interaction with regulatory frameworks.

Expert Interviews: We conducted semi-structured interviews with 15 experts, including compliance officers, AI developers, and regulatory specialists. Participants were selected based on their direct experience implementing or overseeing XAI in compliance contexts.

3.3 Analytical Framework

Our analysis employed a framework that examined XAI implementations across three dimensions:

Technical Implementation: How explainability was achieved technically, including model selection, explanation methods, and validation approaches.

Regulatory Alignment: How implementations addressed specific regulatory requirements and interacted with regulatory bodies.

Organizational Integration: How XAI was integrated into existing compliance processes, governance structures, and stakeholder interactions.

3.4 Limitations

This research has several limitations. First, the rapidly evolving nature of both AI technology and regulatory requirements means that findings may become dated quickly. Second, our case studies were limited by the availability of detailed implementation information, potentially creating selection bias toward more successful or public implementations. Third, the diversity of regulatory contexts makes generalization challenging, though we attempted to identify common patterns across domains.

IV. Analysis and Findings

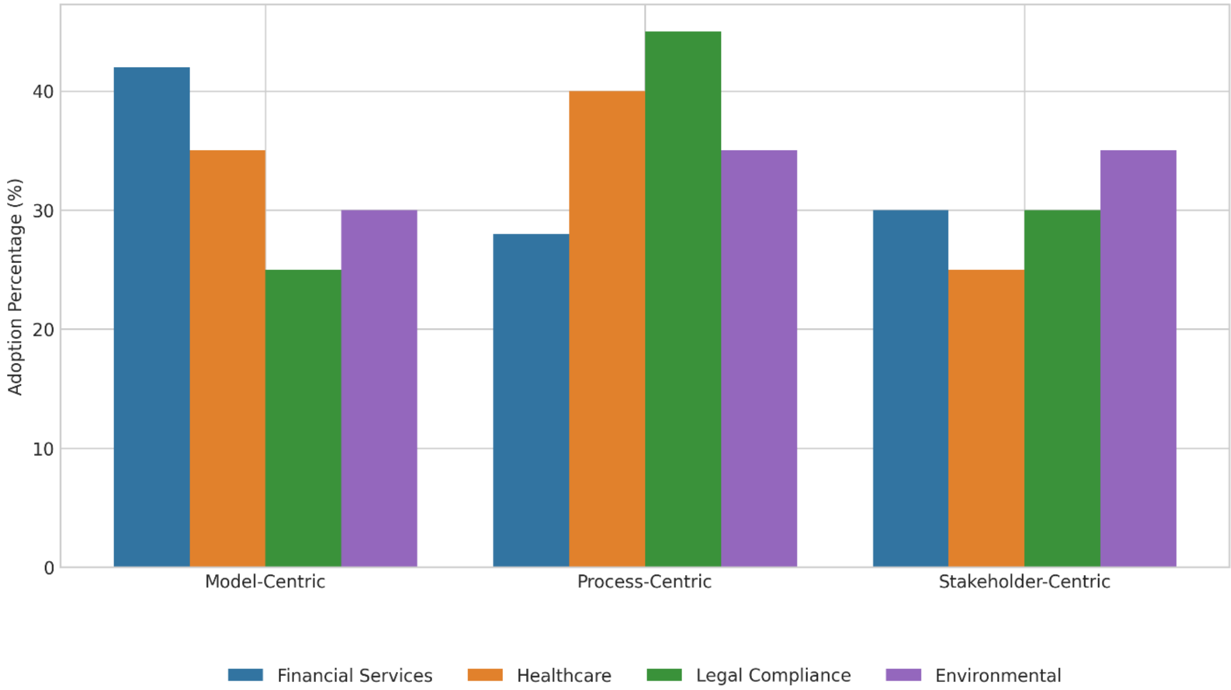
4.1 Current State of XAI Implementation in Compliance Audits

Our research reveals that XAI implementation in compliance audits is still evolving, with significant variation

IJSES
International Journal of Scientific Engineering and Science
Volume 9, Issue 3, pp. 137-144, 2025. ISSN (Online): 2456-7361

across industries and regulatory domains. Organizations are employing three primary approaches:

Figure 3: XAI Implementation Approaches by Industry



Model-Centric Approaches focus on making AI models themselves more interpretable. In compliance contexts, these include using inherently interpretable models like decision trees for high-risk applications, employing model-specific explanation techniques like feature importance measures, and model distillation where complex models are simplified for explanation purposes. These approaches are most prevalent in highly regulated domains like financial services.

Process-Centric Approaches embed explainability throughout the compliance workflow rather than focusing solely on model transparency. These include staged decision systems where AI flags potential issues for human review, comprehensive documentation practices, and continuous monitoring of explanation quality. These approaches recognize that explainability requires integration with governance frameworks and operational processes.

Stakeholder-Centric Approaches tailor explanations to different audiences in the compliance ecosystem. These include layered explanations with varying levels of detail, interactive interfaces allowing stakeholders to explore explanations, and narrative explanations that translate technical details into domain-specific compliance language.

Common challenges in XAI implementation include technical issues (explanation fidelity, performance-explainability trade-offs), organizational barriers (expertise gaps, integration difficulties), and regulatory challenges (evolving requirements, jurisdictional variations, lack of standards).

4.2 Case Studies of XAI in Compliance Contexts

Financial Services Case Study: Anti-Money Laundering Compliance

A global bank implemented an AI-driven anti-money laundering (AML) system using advanced machine learning to identify suspicious transactions. Their XAI implementation combined feature importance measures, rule extraction, and case-based reasoning to generate standardized suspicious activity reports with detailed explanations.

The bank engaged proactively with regulators during development, incorporating feedback on explanation requirements. Regulators initially expressed concerns about the “black box” nature of the models but accepted the implementation based on the comprehensive explanation framework. The bank demonstrated that the AI system with explanations provided more consistent and detailed justifications than previous manual processes.

Key outcomes included a 40% reduction in false positives while increasing identification of genuinely suspicious activities by 23%. The most effective explanations combined technical details with domain-specific narrative that aligned with regulatory language and expectations.

Healthcare Case Study: HIPAA Compliance Monitoring

A healthcare system implemented an AI-driven monitoring system to identify potential violations of the Health Insurance Portability and Accountability Act (HIPAA). The system used natural language processing and pattern recognition to analyze access logs, communications, and document handling.

For each potential violation flagged, the system generated an explanation detailing the specific HIPAA provision potentially violated, the evidence supporting the concern, and the confidence level of the assessment. Explanations were designed to support human investigation, providing sufficient

IJSES
International Journal of Scientific Engineering and Science
Volume 9, Issue 3, pp. 137-144, 2025. ISSN (Online): 2456-7361

detail for compliance officers to quickly evaluate the potential issue.

Healthcare privacy regulators required evidence that the AI system could reliably identify compliance issues without generating excessive false positives. The healthcare system demonstrated that the explainable approach improved compliance by providing more consistent and comprehensive monitoring than manual processes.

The implementation reduced compliance investigation time by 60% while increasing monitoring coverage. The case highlighted the importance of domain-specific explanation frameworks that align with the specific regulatory language of healthcare privacy.

4.3 Regulatory Perspectives on XAI

Regulatory guidance on AI explainability varies significantly across domains and jurisdictions, but several common themes emerge:

Principles-Based vs. Rules-Based Approaches: Many regulators have issued high-level principles for AI transparency without prescribing specific technical approaches, while some frameworks include more specific requirements. Our analysis indicates that principles-based approaches currently predominate, giving organizations flexibility but creating uncertainty about compliance standards.

Documentation Requirements: Regulatory guidance consistently emphasizes documentation as a key element of explainability, including model documentation, decision records, and process documentation. These requirements serve both explainability and auditability purposes.

Human Oversight Emphasis: Regulatory guidance frequently emphasizes human oversight as a complement to technical explainability, including human-in-the-loop requirements, override mechanisms, and expertise requirements.

Regulatory expectations are evolving rapidly, with several key trends:

Increasing Technical Specificity: As regulators develop greater expertise in AI technology, their guidance is becoming more technically specific, suggesting organizations should prepare for more detailed requirements.

Growing Focus on Outcomes and Impacts: Regulatory attention is increasingly focusing on how AI systems affect individuals and society, not just how they function technically.

Convergence Around Key Principles: Despite jurisdictional variations, there is convergence around several key principles: the right of affected individuals to understand decisions, the importance of documentation, the need for human oversight, and the requirement for ongoing monitoring.

V. Discussion and Implications

5.1 Bridging Technical and Regulatory Requirements

Our research reveals a fundamental tension between technical capabilities and regulatory expectations in implementing XAI for compliance. Organizations are addressing this tension through several approaches:

Strategic Model Selection based on risk and regulatory scrutiny, with high-risk compliance decisions using inherently interpretable models while lower-risk functions employ more complex models with post-hoc explanations. Some organizations implement hybrid architectures where complex models identify potential issues, but final determinations use more interpretable models.

Explanation Quality Metrics help ensure explanations meet both technical and regulatory standards, including fidelity measures that assess how accurately explanations represent model decisions, comprehensibility testing with different stakeholders, and regulatory alignment checks against specific requirements.

Continuous Improvement Processes recognize that effective explainability requires ongoing attention, including explanation feedback loops, benchmark comparisons against emerging best practices, and regulatory horizon scanning to anticipate future requirements.

Beyond the performance-explainability balance, organizations must address the broader trade-off between system complexity and transparency through layered transparency approaches, transparency by design principles, and regulatory-technical translation mechanisms that bridge the gap between technical capabilities and regulatory expectations.

5.2 A Framework for XAI Implementation in Compliance Audits

Based on our findings, we propose a framework for implementing XAI in compliance contexts that integrates technical, organizational, and regulatory considerations:

Key Components

1. Regulatory Alignment Strategy: Organizations must develop a clear strategy for aligning XAI approaches with relevant regulatory requirements, including regulatory mapping, compliance prioritization, and regulatory engagement planning.
2. Technical Implementation Architecture: The technical architecture must be designed specifically for compliance contexts, including a model selection framework, explanation generation system, and documentation architecture.
3. Governance and Oversight Mechanisms: Effective XAI requires appropriate governance structures, including clear roles and responsibilities, review and approval processes, and ongoing monitoring mechanisms.
4. Stakeholder Engagement Model: Organizations must engage effectively with diverse stakeholders, mapping their specific explainability needs, developing appropriate explanation delivery channels, and collecting feedback on explanation effectiveness.
5. Continuous Improvement System: XAI implementation must evolve over time, requiring clear performance metrics, structured improvement processes, and knowledge management systems.

Stakeholder-Specific Requirements

Different stakeholders have different explainability requirements in compliance contexts:

Regulatory stakeholders typically require comprehensive documentation, regulatory alignment evidence, audit trails, and verification mechanisms. Explanations should emphasize completeness, accuracy, and alignment with specific regulatory provisions.

Internal compliance teams need risk indicators, exception handling information, process integration, and governance

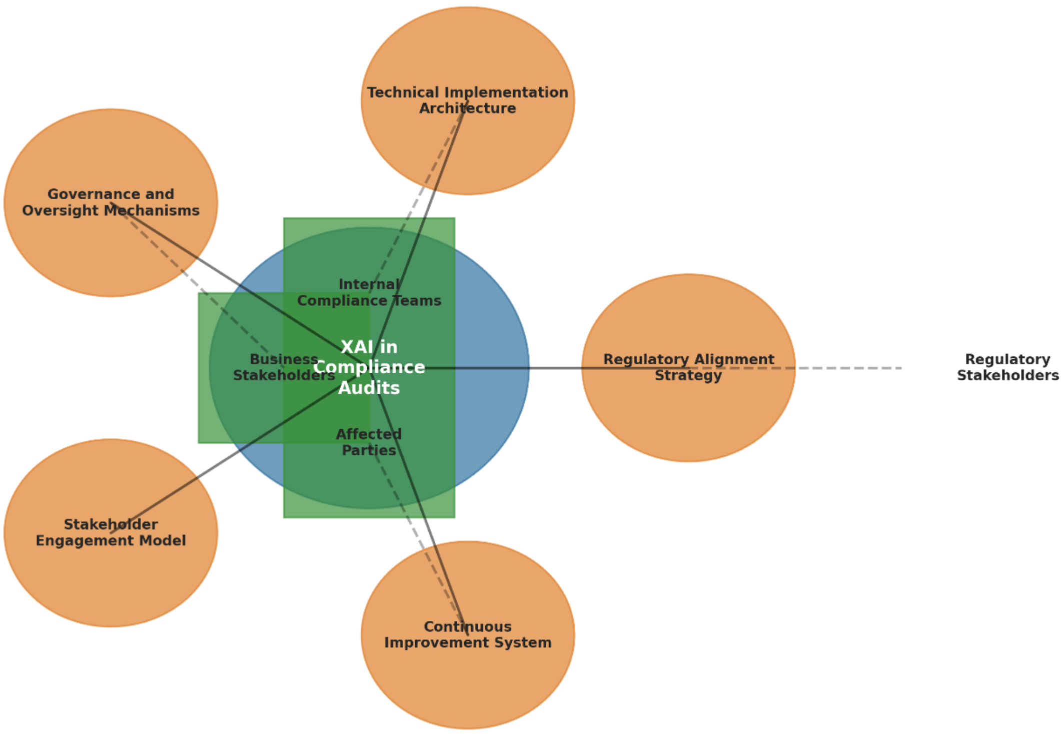
International Journal of Scientific Engineering and Science
Volume 9, Issue 3, pp. 137-144, 2025. ISSN (Online): 2456-7361

support. Explanations should emphasize practical utility for compliance management.

Business stakeholders require business impact context, resource implications, strategic alignment, and competitive context. Explanations should emphasize strategic relevance rather than technical details.

Affected parties need accessible language, actionable insights, comparative context, and recourse options. Explanations should emphasize clarity and relevance rather than technical comprehensiveness.

Figure 4: Framework for XAI Implementation in Compliance Audits



5.3 Recommendations

For Compliance Practitioners

- Conduct a regulatory requirement analysis before selecting XAI approaches
- Adopt a risk-based approach, prioritizing explainability investments based on compliance risk
- Integrate XAI into the broader compliance technology strategy
- Implement explanation by design rather than retrofitting explanations
- Align explanations with existing compliance processes and documentation requirements
- Develop comprehensive documentation frameworks mapped to regulatory requirements
- Create tailored communication approaches for different stakeholders

For Regulators and Policy Makers

- Adopt principles-based approaches focused on outcomes rather than prescribing specific technical methods
- Implement risk-based tiering with graduated requirements based on risk level
- Provide clear guidance and examples to help organizations understand expectations
- Support industry standards development while allowing for innovation
- Consider proportionality in requirements based on organization size and resources
- Participate in international coordination to harmonize cross-jurisdictional approaches

1) *For Technology Developers*

- Incorporate explainability considerations from the earliest stages of system design
- Develop domain-specific explanation tools tailored to compliance contexts

International Journal of Scientific Engineering and Science
Volume 9, Issue 3, pp. 137-144, 2025. ISSN (Online): 2456-7361

- Create explanation validation tools for verifying accuracy and completeness
- Build integration frameworks for connecting explanation capabilities with existing compliance systems
- Implement standardized APIs and exchange formats for explanation interoperability

VI. CONCLUSION

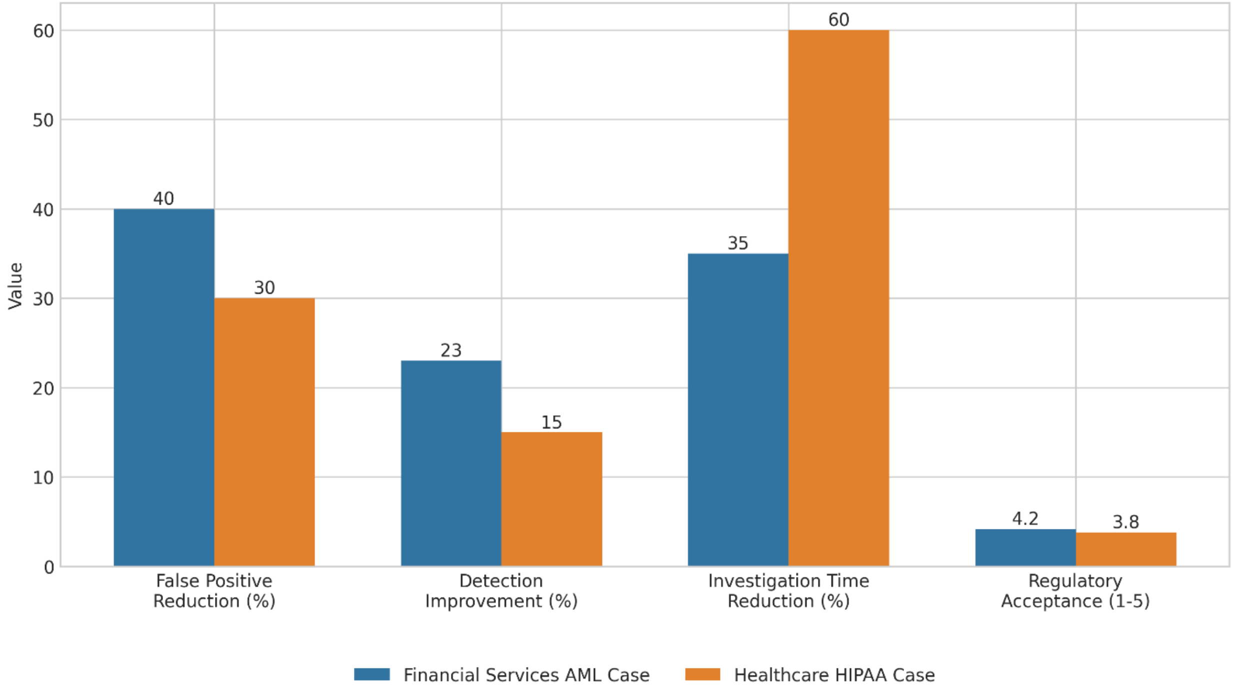
This research has examined the critical role of explainable AI in compliance audits, focusing on how explainability impacts regulatory acceptance and trust in AI-driven compliance decisions. Through analysis of regulatory frameworks, case studies, and implementation approaches, we have demonstrated that explainability is not merely a technical consideration but a fundamental requirement for the successful integration of AI in regulated environments.

Summary of Key Findings

Our research has yielded several important findings regarding XAI in compliance contexts:

The Regulatory Imperative for Explainability has emerged as non-negotiable in regulated environments where AI systems support compliance functions. This requirement stems from fundamental regulatory principles of transparency, accountability, and justifiability that transcend specific technologies. The regulatory landscape continues to evolve, with increasing specificity in explainability requirements as regulators develop greater expertise in AI technology.

Figure 5: XAI Implementation Results from Case Studies



The Technical-Regulatory Gap persists between technical capabilities for explainability and regulatory expectations. Current XAI methods, while advancing rapidly, still face limitations in addressing the full range of regulatory requirements for transparency. Technical explanations often fail to align with regulatory language and concepts, creating implementation challenges for organizations.

Implementation Approaches that successfully bridge this gap include risk-based tiering of explainability requirements, layered explanation approaches tailored to different stakeholder needs, proactive regulatory engagement, and continuous monitoring of explanation quality and effectiveness.

The Multidimensional Nature of Explainability in compliance contexts encompasses technical explainability, regulatory alignment, stakeholder utility, and operational integration. Effective XAI implementation requires addressing all these dimensions, recognizing that different stakeholders have different explainability needs.

Limitations and Future Research

This study has several limitations, including the rapidly evolving nature of both AI technology and regulatory requirements, potential selection bias in case studies, and challenges in generalizing across diverse regulatory contexts.

Future research should focus on developing domain-specific explanation methods tailored to compliance contexts, creating more robust metrics for assessing explanation quality, exploring causal approaches to explainability that better align with regulatory expectations, and investigating effective governance models for managing explainability in compliance functions.

Concluding Remarks

As AI systems increasingly support compliance functions, explainability has emerged as a critical bridge between technological innovation and regulatory requirements. This research has demonstrated that effective explainability in

IJSES

International Journal of Scientific Engineering and Science
Volume 9, Issue 3, pp. 137-144, 2025. ISSN (Online): 2456-7361

compliance contexts requires integrated approaches that address regulatory alignment, stakeholder needs, and organizational implementation.

The path forward involves continued innovation in both technical methods and implementation approaches, with collaboration among technology developers, compliance practitioners, and regulators to develop shared understanding and effective practices. By addressing the challenges identified in this research and building on successful approaches, organizations can harness the power of AI for compliance while maintaining the transparency, accountability, and trust that regulatory frameworks demand.

REFERENCES

[1]. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138-52160.
[2]. Brundage, M., Avin, S., Wang, J., Belfield, H., Krueger, G., Hadfield, G., et al. (2020). Toward trustworthy AI development: Mechanisms for supporting verifiable claims. arXiv preprint arXiv:2004.07213.
[3]. European Commission. (2021). Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
[4]. Felzmann, H., Villaronga, E. F., Lutz, C., & Tamò-Larrieux, A. (2019). Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. Big Data & Society, 6(1).
[5]. Financial Stability Board. (2017). Artificial intelligence and machine learning in financial services: Market developments and financial stability implications.
[6]. Kaminski, M. E. (2019). The right to explanation, explained. Berkeley Technology Law Journal, 34(1), 189-218.
[7]. Kroll, J. A. (2018). The fallacy of inscrutability. Philosophical Transactions of the Royal Society A, 376(2133).
[8]. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (pp. 4765-4774).
[9]. Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1-38.
[10]. Monetary Authority of Singapore. (2019). Principles to promote fairness, ethics, accountability and transparency (FEAT) in the use of artificial intelligence and data analytics in Singapore's financial sector.
[11]. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135-1144).
[12]. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215.
[13]. Thelisson, E., Padh, K., & Celis, L. E. (2017). Regulatory mechanisms and algorithms towards trust in AI/ML. In Proceedings of the IJCAI 2017 Workshop on Explainable Artificial Intelligence (XAI) (pp. 5-9).
[14]. Wachter, S., Mittelstadt, B., & Russell, C. (2018). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. Harvard Journal of Law & Technology, 31(2), 841-887.