

AI-Driven Trade Surveillance Framework for Real-Time Market Manipulation Detection in Financial Institutions

Karthikeyan Thandayutham

Independent Researcher, USA

Abstract

The challenges that affect financial markets are growing ever more difficult since the advanced forms of manipulation cannot be detected properly using traditional surveillance systems. The combination of machine learning and artificial intelligence in trade surveillance is a paradigm shift in how financial institutions can solve the problem of market integrity and regulations. State-of-the-art detection algorithms are able to detect subtle patterns of manipulation that rule-based systems entirely do not consider, whereas explainable AI methods guarantee transparency and acceptance by regulators. As financial institutions continue to struggle with inflated false positive rates, rising compliance expenses, and inefficient use of investigative resources, they face increasing pressure to improve surveillance effectiveness and operational efficiency. Current systems integrate supervised and unsupervised machine learning models with intelligent alert triage functions, human control mechanisms, and automated compliance reporting systems. Its implementation must pay attention to workflow integration, model explainability, governance, and full change management. Evidence from industry reports and academic studies indicates that AI-driven surveillance can significantly reduce false positive rates while improving detection of sophisticated manipulation schemes. Cloud-enabled architectures offer the scalability required in real-time processing as well as meet regulatory requirements of auditability and business continuity. The interplay of increasing costs, advanced threats, and regulatory pressures forms powerful incentives for financial institutions to move towards AI-enhanced surveillance systems. The strategic adoption puts the institutions in a better position to respond to the changing regulatory environment and helps in enhancing the integrity of the market as well as investor confidence. Effective change requires long-term dedication, proper distribution of resources, and an understanding that change is not a technology project but a process that is going to take years.

Keywords: Artificial Intelligence, Trade Surveillance, Market Manipulation Detection, Explainable AI, Regulatory Compliance

1. Introduction

In recent years, financial markets have been transformed drastically. Trade volumes in major exchanges around the world are characterized by high-frequency trading and algorithmic trading. The developments introduced a level of complexity into the operations of the market, which is something the older surveillance can hardly keep up with, as new avenues of manipulation and abuse are introduced. Integrating machine learning with artificial intelligence in the trade surveillance systems is a paradigm shift in the manner in which financial institutions address the issue of market integrity and regulatory compliance. Studies on AI and machine learning in trade surveillance reveal these technologies can identify subtle or sophisticated patterns that traditional rule-based systems completely overlook, while offering better explainability through transparent decision factors that satisfy regulatory demands [1]. This capability becomes more crucial as market manipulation techniques advance in sophistication.

10.48047/jocaaa.2025.34.12.01

The regulatory landscape demands robust surveillance capabilities that work effectively in real-time trading environments. FINRA's 2025 Annual Regulatory Oversight Report emphasizes the ongoing push to modernize surveillance approaches and highlights rising expectations for firms to adopt technology-driven solutions that can catch complex market abuse patterns across various trading venues and asset classes [2]. The financial institutions are fighting the pressure of reducing the false positive rates and increasing the detection rate of cases of actual manipulations. The obstacle goes beyond pure detection to operational efficiency, regulatory reporting, and incorporation of AI-based knowledge into the predetermined compliance processes. This article lays out a comprehensive framework tackling these varied challenges by mixing advanced anomaly detection algorithms with operational governance structures and human oversight mechanisms. The framework stresses practical implementation considerations that help financial institutions transition from traditional rule-based surveillance to intelligent, adaptive systems that meet contemporary regulatory expectations while keeping operational efficiency and audit readiness intact.

2. Problem Context and Current Challenges

Financial institutions in today's trading environments struggle with significant challenges around compliance costs, operational inefficiencies, and traditional surveillance technology limitations. The cost burden tied to financial crime compliance has climbed to unprecedented heights, putting substantial pressure on institutional resources and affecting overall compliance operation quality. Research looking at the high cost of financial crime compliance shows institutions are facing escalating expenses that directly harm customer experience and operational efficiency, with compliance teams battling to manage mounting alert volumes generated by legacy surveillance systems [3]. These cost pressures appear not just in direct spending on compliance personnel and technology infrastructure but also in opportunity costs as resources that could go toward strategic initiatives are consumed investigating false positive alerts instead.

Managing compliance costs intersects with fundamental limitations in detection effectiveness and operational workflow integration. Conventional surveillance systems based on rules utilize fixed limits that flood the system with too many alerts without enough prioritization of risks or situational context. Compliance departments are overwhelmed with alerts that are much larger than they can investigate, causing a backlog and increasing the possibility that truly manipulated cases slip by or might be investigated without their due diligence. The absence of intelligent triage mechanisms implies that there is a similar priority of low-risk and high-risk alerts, and resource allocation is inefficient, and the major cases of serious market abuse are not detected promptly.

Regulatory technology evolution offers potential fixes for these challenges through advanced analytics and automation capabilities. Analysis of RegTech companies and their role in compliance transformation shows emerging technologies let financial institutions boost detection accuracy while cutting operational burden through intelligent automation and workflow optimization [4]. These technical solutions use machine learning algorithms to identify patterns that avoid the conventional rule-based detection, and explainable AI methods make sure that the outputs of algorithms can be verified by compliance experts and presented to regulators in a sufficiently transparent way. Nonetheless, effective implementation requires more than the technology deployment; it requires thorough change management, integration with the current systems, and the creation of hybrid human-AI workflows that preserve the knowledge of the institution and help enhance the capabilities.

10.48047/jocaaa.2025.34.12.01

The broader financial crime context further highlights the urgency for enhanced surveillance capabilities. The PwC Global Economic Crime and Fraud Survey exposes pervasive challenges across the financial services sector, with institutions reporting that fraud and economic crime keep posing substantial threats requiring sophisticated detection and prevention mechanisms [5]. Market manipulation represents a specific category within this broader financial crime landscape, one demanding specialized surveillance approaches tailored to trading environment characteristics. Market manipulation in high-frequency trading situations may occur in a matter of seconds or minutes, unlike in the context of traditional fraud, which can have a long-term nature and necessitate real-time detection features that a traditional compliance system is unable to provide. The combination of cost increases, inefficiencies in operations, technological opportunities, and the ongoing threats triggers the necessity of financial institutions to radically reconsider the use of AI-driven frameworks as a balance between automation and human expertise in terms of trade surveillance.

Challenge Category	Key Issues	Impact on Operations
False Positives	Excessive alert volumes without risk prioritization	Resource drain, investigation backlogs, and delayed detection
Compliance Costs	Escalating expenses for personnel and infrastructure	Reduced strategic investment, opportunity costs
Detection Limitations	Static thresholds miss sophisticated manipulation	Undetected market abuse, regulatory exposure
Workflow Integration	Disconnected systems and manual processes	Operational inefficiency, poor adoption rates
Regulatory Complexity	Multiple jurisdictions with varying requirements	Documentation burden, examination challenges

Table 1: Current Challenges in Trade Surveillance [3, 4]

3. Proposed AI-Driven Trade Surveillance Framework

The suggested framework is a combination of the latest artificial intelligence features and the developed compliance processes to establish a complete surveillance system that addresses both the accuracy of the detection and the efficiency of the functioning. At the foundation sits a sophisticated data ingestion infrastructure that can process high-frequency transaction data, order book dynamics, and cross-market information streams in real-time. This data foundation allows deployment of advanced machine learning models that can spot manipulation patterns through multiple analytical approaches, including supervised learning trained on historical manipulation cases, unsupervised anomaly detection for novel patterns, and ensemble methods that pull together insights from multiple algorithmic perspectives.

The framework treats explainability as a core design principle rather than an afterthought, recognizing that regulatory acceptance and operational adoption hinge critically on transparency. Research on explainable AI for regulatory compliance in financial and healthcare sectors shows that comprehensive approaches to interpretability let organizations satisfy regulatory requirements while keeping high detection accuracy, with transparent decision factors allowing compliance professionals to understand and validate algorithmic outputs [6]. This explainability comes through techniques such as SHAP values that quantify feature-level contributions to individual alert decisions, letting analysts understand exactly which trading characteristics triggered specific alerts and supporting both internal investigation processes and external regulatory inquiries.

10.48047/jocaaa.2025.34.12.01

Integrating artificial intelligence into surveillance operations needs careful attention to how algorithmic outputs interact with human decision-making processes. Advances in machine learning for financial anomaly detection highlight how important hybrid architectures are that tap AI for pattern recognition and prioritization while keeping human judgment for contextual interpretation and final disposition decisions [7]. The structure establishes smart alert triage systems that evaluate the probability of manipulation, possible market effect, trader record, and cross-asset relations to create risk scores that enable compliance offices to devote examination assets to the most high-profile cases. With this risk-based prioritization, the traditional first-in-first-out alert processing is fundamentally changed, which allows allocating resources more efficiently and quickly notice serious manipulation plans.

Operational integration also extends beyond detection algorithms into full automation of workflow and compliance reporting. The framework also produces an audit-ready documentation of every investigation automatically and maintains full traceability between the first detection and final disposition in forms that meet regulatory requirements in a variety of jurisdictions. Connection to the current case management systems will make the alerts generated by AI move seamlessly into the existing investigation processes, and will not disrupt the existing processes but enhance them with intelligent information. The structure inherently incorporates the continuous improvement processes through the feedback loop, which collects the dispositions of analysts and the results of investigations and utilizes the information to refine the model parameters and detection rules in the course of development. This adaptive ability has the effect of making surveillance more efficient as time goes on, as the system gets to learn through successes as well as failed searches, hence forming a continuous improvement surveillance ability that remains efficient despite the manipulation process also improving and the change in the market structure.

Component	Function	Key Features
Data Ingestion	Real-time processing of trading data	High-frequency transactions, order books, market feeds
Anomaly Detection	Pattern recognition across manipulation types	Supervised learning, unsupervised models, ensemble methods
Alert Triage	Risk-based prioritization	Manipulation likelihood scoring, impact assessment
Human Oversight	Expert validation and feedback	Guided workflows, contextual analysis, and disposition capture
Compliance Reporting	Automated documentation generation	Audit-ready records, regulatory filing integration
Continuous Improvement	Adaptive learning mechanisms	Feedback loops, model retraining, and maturity assessment

Table 2: AI-Driven Surveillance Framework Components [5, 6]

4. Implementation Best Practices and Critical Considerations

Effectively establishing AI-based trade surveillance systems requires a careful combination of both technological savvy and operational practicality, and regulatory adequacy. The use of artificial intelligence in the framework of financial market surveillance is associated with specific issues of model validation, explainability, regulatory acceptance, and change management. Research examining AI and machine learning applications in trade surveillance stresses that institutions must prioritize several critical implementation principles to achieve successful outcomes, including integrating AI outputs into existing compliance workflows rather than building parallel systems, maintaining explainable models that compliance professionals and regulators can validate and understand, and establishing clear governance frameworks that define roles, responsibilities, and accountability for surveillance effectiveness [8].

The principle of workflow integration recognizes that even the most sophisticated detection algorithms deliver limited value if they can't be operationalized effectively within established compliance processes. As practice has demonstrated, frameworks that have been successful in integrating AI functionality into an existing operational environment will realize a much higher adoption rate and a much quicker time-to-value than those that demand compliance teams to change to new systems and processes entirely. The integration must also maintain institutional knowledge and established investigation methodologies and enhance them with AI-based insights, prioritization, and automation. This approach minimizes change management resistance and lets compliance professionals focus learning efforts on understanding AI-specific capabilities rather than relearning fundamental investigation processes.

Explainability emerges as a non-negotiable requirement in regulated financial contexts where surveillance decisions must stand up to internal stakeholders, external auditors, and regulatory examiners. Research on explainable AI for regulatory compliance emphasizes that transparent decision-making processes let organizations satisfy increasingly strict regulatory expectations for algorithmic accountability while simultaneously improving operational effectiveness by helping compliance analysts understand and trust AI-generated insights [9]. Explainable AI methods can convert unreadable outputs of algorithms to readable insights that compliance professionals can authenticate, investigate, and submit to regulators with certainty. This openness is especially important when a regulatory examination takes place, and

10.48047/jocaaa.2025.34.12.01

companies should demonstrate that the areas of manipulation were detected by surveillance systems, but also justify why certain patterns caused alerts and others did not.

The governance aspect addresses the organizational framework, policies, and procedures that can guarantee the AI-driven surveillance is working within the right governance and control systems. Good governance creates ownership of model development, validation, and continuous monitoring, sets up the process of escalation of alert investigation and regulatory reporting decisions, and retains the detailed documentation of surveillance methodology and performance measures. Best practices in implementation emphasize that although AI can be used to automate pattern recognition and prioritize alerts, any final decisions on regulatory filings and enforcement referrals must remain in human hands to hold them accountable and add contextual variables that can be poorly forecasted by models. Organizations must also avoid the temptation to pursue the highest level of automation as an end goal, but should maximize the overall effectiveness and efficiency of integrated human-AI systems. The most effective applications understand that reducing false positives may require some strategic human intervention at the critical decision-making points and that hybrid solutions involving the use of AI to do extensive monitoring with human judges to provide fine-tuning judgments tend to outperform those that rely on either a fully manual or entirely automated solution.

Practice Area	Recommended Actions	Expected Outcomes
Workflow Integration	Embed AI into existing compliance processes	Higher adoption rates, minimal disruption
Explainability	Implement SHAP, LIME techniques	Regulatory acceptance, analyst confidence
Governance	Define roles, responsibilities, and accountability	Clear oversight, comprehensive documentation
Change Management	Prioritize training, communication, and phased rollout	Staff proficiency, cultural alignment
Data Quality	Establish monitoring, validation, and lineage tracking	Model reliability, consistent performance
Human-AI Balance	Preserve human control for critical decisions	Optimal effectiveness, regulatory compliance

Table 3: Implementation Best Practices [7, 8]

5. Evidence from Literature and Broader Implications

While this paper does not present institution-specific experimental deployments, extensive evidence from published studies, regulatory reports, and industry analyses demonstrates the benefits of integrating AI into trade surveillance environments. Across the literature, machine learning techniques consistently outperform traditional rule-based systems in detecting subtle, nonlinear, and high-frequency manipulation behaviors that static thresholds often miss [7], [10]. These works show that supervised models effectively capture known abuse patterns, while unsupervised algorithms identify previously unseen anomalies that resemble emerging manipulation schemes, providing broader surveillance coverage.

Industry and regulatory analyses also highlight improvements in operational efficiency when AI techniques are used to support alert triage and investigative workflows. Reports from FINRA, PwC, and RegTech surveys indicate that AI-driven prioritization helps compliance teams focus on higher-risk alerts

10.48047/jocaaa.2025.34.12.01

by reducing redundant or low-quality cases without compromising regulatory expectations [2], [5]. Although quantitative impacts vary across studies, the evidence consistently demonstrates improvements in analyst productivity, reduction in manual overhead, and enhanced investigative accuracy.

From a market integrity perspective, academic research suggests that stronger surveillance capabilities contribute to better price discovery, reduced informational asymmetry, and lower prevalence of abusive practices. While these systemic effects are influenced by many factors, enhanced detection capabilities play a contributing role in discouraging manipulation attempts and supporting fair market behavior.

Regulators across multiple jurisdictions increasingly recognize the value of AI-enabled analytics, provided that institutions maintain strong governance, explainability, and human oversight. The expectation is shifting away from purely rule-based monitoring toward hybrid architectures that combine data-driven detection with transparent justification of algorithmic outcomes. This regulatory evolution suggests that AI-driven frameworks will become a foundational component of modern surveillance programs.

Looking ahead, emerging innovations in real-time data engineering, graph-based anomaly detection, cross-market correlation analytics, and advanced explainability methods are likely to expand the capabilities of surveillance platforms. These technological advancements will further support scalability, accuracy, and adaptability, positioning AI-driven surveillance as an essential tool for navigating increasingly complex global markets.

Impact Area	Illustrative Benefits	Significance
Detection Accuracy	Reduced false positives, improved true positives	Efficient resource allocation, better threat identification
Operational Efficiency	Lower investigation time, reduced staffing needs	Direct cost savings, enhanced productivity
Regulatory Relations	Improved examination outcomes, fewer findings	Collaborative relationships, reduced scrutiny
Market Integrity	Enhanced deterrence, increased investor confidence	Systemic stability, fair price discovery
Economic Benefits	Lower compliance burden, avoided penalties	Positive return on investment, competitive advantage
Technology Trajectory	Real-time capabilities, cross-border integration	Future readiness, adaptive capabilities

Table 4: Evidence and Broader Implications [9, 10]

Conclusion

The growing complexity and speed of modern financial markets demand surveillance capabilities that extend beyond the limitations of traditional rule-based systems. AI-driven approaches—supported by machine learning, explainable AI, and scalable cloud architectures—provide powerful tools for detecting sophisticated manipulation patterns, reducing false positives, and improving investigative efficiency. These technologies have matured to the point where they can operate within regulated environments while meeting requirements for auditability, validation, and transparency.

The framework presented in this paper integrates real-time data processing, anomaly detection models, explainability techniques, risk-based alert triage, human–AI hybrid investigations, automated reporting, and governance mechanisms into a unified surveillance architecture. Rather than replacing human expertise, the proposed approach enhances analysts’ ability to focus on high-value investigations and ensures that surveillance decisions remain transparent and defensible.

Adopting AI-enhanced surveillance is not a single-step technology deployment but a multi-year transformation requiring strong governance, staff training, workflow integration, and continuous model refinement. Institutions that embrace these capabilities will be better prepared to meet evolving regulatory expectations, reduce operational burden, and strengthen overall market integrity. As financial markets continue to evolve, AI-driven surveillance frameworks will play a critical role in ensuring fair trading environments and maintaining investor confidence.

References

- [1] Trapets, "AI and machine learning in trade surveillance: a 2025 guide," 2025. [Online]. Available: <https://www.trapets.com/resources/blog/ai-machine-learning-trade-surveillance#:~:text=The%20opportunity%20AI%20brings&text=Advanced%20detection%20of%20subtle%20or,explainability%20with%20transparent%20decision%20factors>
- [2] Financial Industry Regulatory Authority (FINRA), "2025 FINRA Annual Regulatory Oversight Report," 2025. [Online]. Available: <https://www.finra.org/sites/default/files/2025-01/2025-annual-regulatory-oversight-report.pdf>
- [3] Banking Frontiers, "High cost of financial crime compliance affects quality CX," 2024. [Online]. Available: <https://bankingfrontiers.com/high-cost-of-financial-crime-compliance-affects-quality-cx/>
- [4] Deloitte, "RegTech Universe 2024". [Online]. Available: <https://www.deloitte.com/lu/en/Industries/technology/analysis/regtech-companies-compliance.html>
- [5] PwC, "PwC's Global Economic Crime and Fraud Survey 2022," 2022. [Online]. Available: <https://www.pwc.com/gx/en/forensics/gecsm-2022/PwC-Global-Economic-Crime-and-Fraud-Survey-2022.pdf>
- [6] Nikhil Gupta, "Explainable AI for Regulatory Compliance in Financial and Healthcare Sectors: A comprehensive review," IJAEM, 2025. [Online]. Available: https://ijaem.net/issue_dcp/Explainable%20AI%20for%20Regulatory%20Compliance%20in%20Financial%20and%20Healthcare%20Sectors%20%20A%20comprehensive%20review.pdf
- [7] Mudita Kohli and Indu Chhabra, "A comprehensive survey on techniques, challenges, evaluation metrics, and applications of deep learning models for anomaly detection," SN, 2025. [Online]. Available: <https://link.springer.com/article/10.1007/s42452-025-07312-7>
- [8] Mostafa Goudarzi and Flavio Bazzana, "Identification of high-frequency trading: A machine learning approach," Research in International Business and Finance, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0275531923002040>
- [9] Ziqing Xia et al., "A systematic review on human-AI hybrid systems and human factors in air traffic management," IJAEM, 2025. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/09544828.2025.2509056?af=R>
- [10] Qiuliuyang Bao et al., "A Deep Learning Approach to Anomaly Detection in High-Frequency Trading Data," AEFR. [Online]. Available: <https://www.aefr.eu/fr/evnement/487/technological-frontiers-in-finance-a-perspective-on-financial-markets-stability-related-to-ai-quantum-computing-and-cyber-developments#:~:text=The%20rapid%20advancement%20of%20artificial,approaches%20to%20cyber%20Orisk%20management>