Review of Futures Markets (RFM)
ISSN: 1933-7116
Volume 34 Issue 2 (2026)
Website: https://fra.st

Forensic AI in Derivatives Markets: Evidentiary Standards for AI-Based Detection Tools in Futures Market Manipulation Enforcement

Sakshi Rewaria[1], Shringarika[2], Harmandeep Kaur[3], Swati Pal[4]

[1] Lecturer, Indian Institute of Management, Rohtak, Haryana, India
Orcid: 0009-0004-8066-660X; Email: advocate.sakshirewaria@gmail.com

[2] Assistant Professor, IPEM Law Academy, Affiliated to CCS University, Ghaziabad, Uttar Pradesh, India

[3] Lecturer, Indian Institute of Management, Rohtak, Haryana, India
Orcid Id : 0009-0000-2821-5909

[4]Assistant professor, IPEM Law Academy, Ghaziabad, Uttar Pradesh, India

Received: 14/01/2026 Accepted:10/04/2026 Published: 22/08/2026

Abstract

In the era of artificial intelligence (AI), financial market regulators must utilise AI to identify today's more advanced methods of manipulation in futures and derivatives markets. New machine-learning algorithms and anomaly detection methods, graph analytics, and synthetic-content detection systems can now rapidly and accurately detect suspicious trading activity like spoofing, layering, wash trading, and AI-generated manipulation in ways that were previously impossible. The trend of using AI-generated forensic outputs in investigations of market abuse is gaining momentum among regulatory bodies, such as the Commodity Futures Trading Commission (CFTC), the European Securities and Markets Authority (ESMA), and the Securities and Exchange Board of India (SEBI). Even with these technological advances, there is considerable uncertainty in the law regarding the evidentiary value of algorithmic findings. AI systems often provide probabilistic results without clear explanations of the reasoning behind them, which has led to them being called the "black box". This presents difficulties regarding the explainability and reproducibility of the AI model, as well as procedural fairness and the ability to meaningfully contest AI-generated evidence in enforcement actions. This paper is a critical analysis of the novel use of AI-based forensic evidence in enforcement of derivatives market manipulation, including a discussion of the CFTC's requirements for explainability, and considers the effectiveness of current regulatory frameworks, including those in the EU (ESMA) and India (SEBI). The study employs a qualitative doctrinal and comparative legal research approach to examine current regulatory guidance, international principles on AI governance, and recent scholarly work to uncover gaps in evidence. The results indicate that, in the eyes of regulators, explainable AI and audit trails are gaining in significance, but current regulations lack a uniform disclosure requirement regarding algorithmic transparency, model validation, error rates, or independent testing. To overcome these limitations, the paper proposes a framework for Evidentiary Explainability (EEF) with

five pillars: algorithmic transparency, reliability assessment, auditability, independent validation, and procedural fairness. The aim of the proposed framework is to enhance the credibility, admissibility, and accountability of forensic evidence produced by artificial intelligence, whilst ensuring effective market oversight and respecting proper due process in the still-evolving field of financial markets.
Keywords: Artificial Intelligence (AI); Derivatives Markets; Futures Market Manipulation; Explainable Artificial Intelligence (XAI); AI-Generated Forensic Evidence; Algorithmic Surveillance; Market Abuse Detection; Financial Regulation; Evidentiary Standards; CFTC.

1. Introduction

1.1 Background

In the last 20 years, financial markets around the world have undergone an unprecedented technology revolution. As a result of the quick growth of electronic trading platforms, algorithmic trading strategies, high-frequency trading (HFT), and automated order execution systems, futures and derivatives trading has fundamentally changed. Once a process involving human action, this is now done in microseconds by complex computer algorithms that can process vast amounts of market information in real time. These technological developments have enhanced the efficiency, liquidity, and price discovery of the market, but have also introduced ever more sophisticated methods of market manipulation that are far more challenging to identify with traditional market surveillance.

Today, with the advent of automated trading algorithms that can execute thousands of orders in mere milliseconds, market abuse is often a more subtle, automated form of manipulation, different from past methods based on manual trading. Today, however, practices such as spoofing, cross-market manipulation, wash trading, momentum ignition, quote stuffing, and intelligent layering are all carried out by intelligent trading systems that evolve in line with prevailing market conditions. More recently, however, the advent of generative AI has added to the risks via the potential use of synthetic news, fake financial reports, AI-generated deceptive executive statements, and AI-generated misinformation that can impact futures prices before regulators can catch up. These have greatly enlarged the scope and complexity of financial market abuse.

Traditional market surveillance systems were mostly rule-based, comprising some predefined thresholds and manually built detection rules. Such systems are efficient at flagging obvious mistakes, but they fail to capture intricate behaviour where several trading accounts are involved, there is cross-market coordination, complex trading algorithms are in play or AI is used to manipulate trades. As a result, the relevant regulators have increasingly turned to AI as a key tool in their market surveillance efforts. Today, millions of trading events are analysed every day using machine learning algorithms, where anomalies and patterns that would otherwise not have been noticed by humans are identified. AI can expedite the identification of suspicious activity more than ever by bringing statistical learning, natural language processing, network analytics, and predictive modelling together.

The technology has been embraced by regulatory agencies across key financial jurisdictions. The U.S. Commodity Futures Trading Commission (CFTC) has adopted new and sophisticated technology to detect fraudulent trading activity in derivatives markets (CFTC, 2024). Likewise, the European Securities and Markets Authority (ESMA) has been asking for the use of innovative analytics to enhance market abuse detection under the Market Abuse Regulation (MAR) (European Securities and Markets Authority, 2024), and the Securities and Exchange Board of India (SEBI) has increased investments in AI-powered surveillance to better regulate the rapidly expanding Indian securities and derivatives markets. Clacks/clearing corporations and exchanges are also increasingly using AI-driven monitoring systems that can continuously monitor order books, transaction history, communication channels and news sources.

The use of AI has significantly enhanced enforcement capabilities. Rather than waiting for a retrospective investigation, regulators can now detect suspicious trading patterns near real time. AI-driven systems can be constantly connected to the market environment and automatically flag trades with an unusual pattern as potentially manipulative, identify hidden ties between traders, and alert traders to the risk of manipulation in specific trading strategies. These systems can drastically decrease the investigative burden and help police and law enforcement allocate resources more effectively.

But with all this technological development has come a new legal and evidentiary issue. Outputs created by AI are being used more and more to become a key ingredient in enforcement evidence. A surveillance system can show a trader, with a 94% probability, that they were involved in trading manipulation; it can reveal information about trading patterns and show if a news article is a fake, created by artificial intelligence. These outputs are increasingly a factor in enforcement decisions, administrative proceedings, settlement negotiations and, perhaps, judicial decisions.

But not every AI can explain the process it went through to draw its conclusions in the way that an expert who normally would provide testimony can. These complex deep learning architectures can be likened to a "black box," making it very hard for investigators, defendants, judges, or administrative tribunals to determine exactly how specific variables affected the final determination. Thus, the evidentiary value of AI-produced forensic products is hard to assess with traditional criteria.

1.2 Research Problem

As more and more people rely on AI-powered surveillance, there is a major disconnect between what technology can offer and its legislation. The established rules of evidence for human investigators and expert witnesses, which are based largely on the principles of analytical reasoning, were formulated for that type of investigation and evidence. The assumptions come into question when AI outputs are forensic, as when an algorithm makes a decision, there is not necessarily the same level of transparency.

Today, machine learning models, especially deep neural networks and ensemble learning algorithms, produce predictions relying on very complex mathematical relationships between thousands or even millions of parameters. While these systems can be very accurate in their predictions, the reasoning processes that they use are not readily understood. In practical terms,

this means that an enforcement authority can base its decision on an algorithmic determination that there is a high probability of "spoofing" or "market manipulation," but not have a clear description of why the model gave the conclusion.

This is not easy to explain, which raises a number of legal issues. First, AI evidence might be inaccessible in enforcement proceedings, and therefore, the analytical pathway would not be available to strongly contest evidence created by AI. Second, courts and administrative tribunals could be challenged in assessing the reliability, relevance, and methodological validity of the algorithmic outputs themselves in the context of the evidentiary requirements for the same. Third, self-regulation can be challenging for regulators, especially when it comes to proving that AI systems are consistently, fairly, and free of unwarranted bias or error.

As AI systems become more than just aids for investigators, they become more of an expert witness; the issue becomes even more critical. While there is no universal consensus on standards to govern the transparency, validation, reporting of uncertainty, and independent verification of algorithms, probability scores, anomaly rankings, behavioural classifications, and predictive risk assessments are becoming factors in enforcement decision-making. Although these concerns are recognised in recent regulatory initiatives, these tend to be less specific on what type of tangible evidence is required for the evidence produced by AI to be deemed sufficiently reliable.

The CFTC's new guidance on explainable AI is an important move by the regulator (CFTC, 2024), but there is considerable uncertainty about the degree of explainability, documentation, auditability, and disclosure that should accompany AI-generated forensic products before their use in enforcement proceedings. The challenge of governing AI also exists in European and Indian regulatory systems, which are still developing and adapting as AI surveillance technologies rapidly grow.

1.3 Research Objectives

The aim of this research is to investigate the evidentiary considerations for AI-generated forensic products in enforcement actions against market manipulation in derivatives. Specifically, to determine how to:

1. Review the increasing importance of Artificial Intelligence and how it is used in identifying manipulation in futures and derivatives markets.
2. Understand the legal and evidential issues with AI forensic products for regulatory action.
3. Critically discuss the regulatory measures that have been taken by CFTC, ESMA and SEBI in terms of explainability, transparency and governance of AI.
4. Identify the current issues and shortcomings of disclosure obligations for algorithms, including issues of transparency, model validation, auditability and error reporting.
5. Propose a practical Evidentiary Explainability Framework (EEF) to enhance the soundness of AI-generated evidence in financial market enforcement, in terms of admissibility, credibility and fairness.

1.4 Research Question

The main research question of the study is:

What are the evidentiary and disclosure requirements that regulators and adjudicators should put forth in connection with AI-generated, as opposed to manual, forensic products (such as spoofing detection, synthetic-news detection and algorithmic manipulation flags) deployed in futures and derivatives market enforcement actions, and how well do existing explainability requirements under the CFTC meet these standards?

In order to answer this question, the paper also addresses the following sub-issues:

- What kind of forensic systems do financial regulators use today that are based on AI?
- What are the legal issues with the use of AI-generated evidence?
- What gaps are there in the existing guidance on explainability and disclosure?
- What are the minimum evidentiary standards that should be used for financial enforcement using AI?

1.5 Significance of the Study

While many studies have examined the use of artificial intelligence in algorithmic trading, financial surveillance, and market abuse detection, there has been limited focus on the admissibility and evidentiary value of AI forensic products. The majority of the literature assesses the technical performance of machine learning models in terms of prediction accuracy, computational efficiency and/or detection ability. Few studies consider the extent to which AI systems' results comply with the procedural safeguards expected for forensic evidence in the context of regulatory enforcement.

The purpose of this paper is to close this gap and combine financial regulation, administrative law, digital evidence, and the explainability of artificial intelligence. AI-generated content is not only considered a form of surveillance but a new type of “evidence” that must be treated as such, with its own evidentiary requirements. The paper adds to the growing body of research on governance of AI in financial markets.

The main contribution of this research consists of a new Evidentiary Explainability Framework (EEF), which sets minimum standards for the application of AI-generated forensic evidence in enforcement in derivatives markets. It has five interrelated pillars: algorithmic transparency, reliability assessment, auditability, independent validation and procedural fairness. These pillars aim to guarantee a technologically effective and legally defensible AI-assisted enforcement.

Moreover, it provides some practical policy suggestions for regulators, adjudicators, financial institutions and technology developers. The study's recommendations include introducing standardised disclosure requirements and evidence-protection measures to enhance regulators' efficiency in using AI tools and to foster transparency and trust in the enforcement process.

2. Literature Review

The use of Artificial Intelligence (AI) in the surveillance of financial markets has sparked significant interest within the academic community across finance, law, computer science, and regulatory governance. Existing literature shows that AI can significantly improve the speed and accuracy of market abuse detection, especially in highly automated futures and derivatives

markets. Researchers have also examined the technical aspects of AI surveillance systems' capabilities, but much less work has been dedicated to assessing the legal veracity and admissibility of AI outputs for forensic purposes. The literature reviewed in this section is divided into four general topics: AI applications in detecting market manipulation, explainable artificial intelligence (XAI), regulatory measures on AI-enabled enforcement, and a lack of literature.

Artificial Intelligence in Detecting Market Manipulation

2.1 Artificial Intelligence in Detecting Market Manipulation

The evolution of algorithmic and high-frequency trading has changed the way that trading is done in the market. Traditional detection approaches, based mostly on manually created rules and statistical thresholds, are less effective at detecting advanced manipulation carried out at machine speed. This has led researchers to present many AI-driven solutions that can detect intricate trading patterns by recognising them with automated pattern recognition and predictive analytics.

The use of machine learning methods has been proven to be very effective in identifying manipulation tactics like spoofing, wash trading, insider trading, momentum ignition, quote stuffing and layering. Historical enforcement data has been used to train supervised learning algorithms such as Random Forests, Support Vector Machines (SVM) (Aggarwal & Wu, 2006; Comerton-Forde & Putniņš, 2011), Gradient Boosting and Neural Networks and high predictive accuracy has been achieved. These models learn the behavioural attributes of the classes of behaviour they have been trained on and use this knowledge to apply to new trading data.

Labelled enforcement data are also scarce, and thus unsupervised learning methods are becoming more popular. Clustering algorithms, anomaly detection, and density-based learning techniques detect unusual trading without the need for a set of labels for manipulation. These can be especially helpful in identifying new forms of manipulation, not seen before in enforcement cases.

Deep learning architectures have also enhanced the capabilities of surveillance, more recently. Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) models, Graph Neural Networks (GNNs) (Aggarwal & Wu, 2006; Comerton-Forde & Putniņš, 2011), and Transformer-based architectures are used to analyse sequential market data, the dynamics of the order book, and relationships among market participants. These systems can identify coordinated trading activities across various exchanges and study transaction timing, order cancellations, communication records, and market sentiment.

Another significant surveillance tool is Natural Language Processing (NLP). AI systems are now being used more and more to analyse corporate disclosures, regulatory documents, news articles, analyst reports, and social media posts to detect false information that can impact derivatives prices. The rise of generative AI has added to regulators' concerns as synthetic news, deepfake executive announcements and AI-driven financial misinformation can quickly impact trading actions before traditional verification processes can react.

In conclusion, the literature shows that AI has a significant impact on the efficiency, scalability and predictive power of financial market surveillance. However, numerous studies tend to focus on detection accuracy, and comparatively little attention is paid to evidence-based standards, procedural fairness, and legal transparency.

2.2 EAI and Digital Forensic Evidence

With the rise in complexity of AI models, there has been a growing number of models that outperform traditional statistical models, but this has created a kind of "black box problem (Rudin, 2019; Adadi & Berrada, 2018). Deep neural networks are commonly able to make very precise predictions, but they don't explain how certain input variables led to the final result. This lack of interpretability has emerged as one of the biggest challenges towards legal acceptance of AI-generated evidence.

Explainable Artificial Intelligence (XAI) has consequently become a field of great interest for studying to enhance the transparency of algorithms while retaining as much predictive power as possible. The XAI project aims to make algorithm decisions comprehensible and to explain the reasons for choosing features, paths, confidence, and how the algorithm behaved.

There are several explanation techniques that have been developed for this purpose. Local Interpretable Model-Agnostic Explanations (LIME) (Ribeiro et al., 2016) create simplified local models to offer explanations for individual predictions, and Shapley Additive exPlanations (SHAP) (Lundberg & Lee, 2017) measure the differential impact of each feature on a model's output. The counterfactual explanations also increase transparency by highlighting minimal changes that would lead to different classification results. Taken together, these techniques enhance understanding and trust in AI-driven decision-making.

In financial regulation, explainability is gaining importance due to its potential impact on legal rights, financial sanctions, and market reputation. Regulatory AI is not used in a commercial recommendation setting (where the user doesn't have a "procedural right" to review, rebut, argue, or reverse the recommendation) but rather in a setting where one or more parties are adversely affected by a recommendation and have such a right.

Digital forensic literature also highlights the importance of the reliability, repeatability, traceability, authenticity and methodological transparency of the evidence. In traditional forensic science, the documentation of all evidence collection, preservation, analysis, and reporting steps is maintained through extensive audit trails. Similar quality is yet to be achieved in AI-generated forensic products.

Explainability alone, on the other hand, is becoming an increasingly frequently voiced argument for a lack of evidentiary reliability. However, a transparent algorithm can give biased or incorrect results in cases where the training set is too small, there is insufficient or poor algorithm validation, or the deployment environment differs from the training environment. Therefore, there is a call to integrate explainability with other governance tools such as independent validation, uncertainty reporting, monitoring and auditing of algorithms.

The discussions suggest that AI-generated forensic evidence should be assessed not only on predictive accuracy but also on other legal considerations, such as transparency, accountability, fairness, and procedural justice. Real enforcement precedent illustrates why the evidentiary

standards discussed above are not merely theoretical. In *United States v. Coscia*, the CFTC and the U.S. Department of Justice brought the first prosecution under the Dodd-Frank Act's anti-spoofing provision, resulting in a 2015 jury conviction of Michael Coscia, principal of Panther Energy Trading LLC, on six counts of spoofing and six counts of commodities fraud for placing large orders he intended to cancel before execution in order to move futures prices (CFTC, 2013; U.S. Department of Justice, 2016). Similarly, the CFTC's 2015 enforcement action against Navinder Singh Sarao and Nav Sarao Futures Limited alleged that his layering and spoofing conduct in E-mini S&P 500 futures contributed to the market conditions behind the May 2010 "Flash Crash," a matter that concluded with Sarao's guilty plea to wire fraud and spoofing (CFTC, 2015; U.S. Department of Justice, 2025). Both cases were detected and proven using conventional order-book and trade-reconstruction evidence rather than AI-generated analysis, which underscores the central problem this paper addresses: as regulators increasingly substitute machine-learning-based anomaly detection for the manual reconstruction methods used in Coscia and Sarao, the evidentiary record supporting a spoofing or layering finding will need to satisfy the same standards of reliability and contestability that manual reconstruction evidence already meets, without yet having a settled body of law or regulatory guidance specifying how an AI-generated finding should be validated, disclosed, or challenged to the same effect.

2.3 Ethical decision-making is crucial for the future of AI

Financial regulators have begun integrating AI governance principles into market surveillance and are recognising both opportunities and risks posed by AI tools in enforcing regulations. But the level of evidence is still not harmonised globally, and there is no international standard for evidence.

The Commodity Futures Trading Commission (CFTC) has become one of the regulators actively looking into the regulation of AI in the derivatives sector. The CFTC has recently placed a greater focus on explainable AI, on AI model governance, on documenting the AI models, on risk management and audit trails of AI systems in compliance. These efforts recognize that outputs created by AI tools need to be transparent enough to enable regulatory accountability. However, the current guidance offers only general guidelines, not thresholds of evidence of acceptable levels of explainability, validation and disclosure(CFTC, 2024; CFTC, 2025).

In Europe, the European Securities and Markets Authority (ESMA) has been working on a broader range of financial regulatory initiatives to foster the responsible use of AI, such as the Market Abuse Regulation (MAR), the Digital Operational Resilience Act (DORA) and the European Union's AI Act. Transparency, accountability, human oversight and risk-based governance are key features of the regulatory philosophy in Europe. The high-risk applications of AI are expected to meet enhanced documentation, monitoring and conformity assessment requirements. But ESMA has not yet established any detailed evidentiary standards that are specifically applicable in the context of market abuse investigations using AI-generated forensic evidence (European Parliament and Council of the European Union, 2024; ESMA, 2024).

Similarly, SEBI in India has its own initiatives to implement AI-driven surveillance systems in securities exchanges and market infrastructure institutions.SEBI has also hastened the rollout of AI-based surveillance tools in securities exchanges and market infrastructure institutions in India. There are a number of ways in which AI is now being used to track trading activities, notice abnormal trading patterns, detect insider trading and enhance the efficiency of regulators. To promote responsible governance of AI, SEBI has issued consultation papers on the use of AI and machine learning in the financial markets. Even with these advances, there are relatively few details about the procedural standards required for algorithmic explainability, disclosure of evidence, independent validation, or procedural safeguards.

Other international bodies, such as the International Organization of Securities Commissions (IOSCO) and the Organisation for Economic Co-operation and Development (OECD), have also expressed a focus on AI principles of transparency, accountability, fairness and human oversight. However, these principles are generally policy-oriented and offer scant guidance on how AI-generated forensic products can be used as evidence in enforcement proceedings(IOSCO, 2021; OECD, 2019).

As a result, while regulators are starting to appreciate the need for trustworthy AI, there is a lack of cohesive governance, and most current rules focus on the ethical use of AI, not the admissibility of evidence.

2.4 Research Gap

Existing research shows significant advances in developing AI models to detect advanced forms of derivatives market manipulation. Existing literature highlights significant advances in developing AI models to detect advanced derivatives market manipulation. Major progress is also being made in the field of explainable AI (XAI), algorithmic governance and financial market surveillance. However, an aspect of financial regulation, digital evidence and administrative law is lacking altogether.

Existing studies mostly focus on technical evaluations of AI, including prediction accuracy, computational speed, and detection efficiency. There are very few studies that explore the question of whether the forensic outputs of AI meet the rigour and standards of admissible evidence as established by traditional rules. The extent to which such questions can be addressed - algorithmic transparency, algorithmic reproducibility, uncertainty reporting, independent validation and procedural fairness - remains an inadequately explored aspect of derivatives market enforcement.

Additionally, existing regulations (such as CFTC, ESMA, SEBI) promote explainability and good governance of AI but do not go so far as to define clear disclosure mandates. There is currently no overarching guidance on what information, if any, regulators are required to provide when using algorithmic outputs in enforcement proceedings, including how the algorithm was designed, model validation, confidence levels, audit trails, and error rates.

This work seeks to address this by shifting from the perspective that AI is a surveillance tool to a lens focused on AI as forensic evidence. It suggests an Evidentiary Explainability Framework (EEF) grounded in foundational concepts of explainability, reliability, auditability,

independent validation, and procedural fairness, and is particularly tailored to the context of AI-driven enforcement of market manipulation in the derivatives market.

3. Methodology

3.1 Research Design

In this study, the researcher uses qualitative doctrinal legal research with comparative regulatory analysis as the method to review the evidentiary standards that can be used in enforcing the futures and derivatives market against market manipulation using AI-generated forensic outcomes. The study does not claim to be empirical, but rather conceptual, focusing on assessing current legal and regulatory frameworks rather than predicting the performance of particular machine learning models.

This research adopts a doctrinal methodology because it involves the analysis of the regulatory guidance, legal provisions, policy documents, administrative considerations and academic research on artificial intelligence, financial regulation, and digital evidence. The comparative analysis also helps identify the similarities and differences between the leading regulatory jurisdictions, in particular, the United States, the European Union, and India.

Finally, the study aims to determine whether the current explainability requirements are sufficient to ensure the admissibility, reliability, and procedural fairness of AI-generated forensic evidence in market-manipulation proceedings, and to propose a common evidentiary framework to address gaps in market-manipulation legislation.

3.2 Data Sources

This study is based only on secondary data, using only authoritative legal, regulatory, and academic sources. Regulatory documents were considered because they provide the official stance of financial supervisory authorities on AI governance and market surveillance. Academic works offer a theoretical background for the analysis of explainability, digital evidence and the enforcement of regulations with the support of AI.

The following are the main sources studied:

- Governance docs and guidance from the Commodity Futures Trading Commission (CFTC).
- Reports and market abuse regulatory publications from the European Securities and Markets Authority (ESMA).
- Consultation papers and papers relating to surveillance by the Securities and Exchange Board of India (SEBI).
- The European Union Artificial Intelligence Act (EU AI Act) was recently passed.
- A regulation known as Market Abuse Regulation (MAR). Copyright © 2019, The Bank of England (BoE).
- Bank of England & Financial Conduct Authority. (2019). *Machine learning in UK financial services.*
- Artificial intelligence and market integrity are topics of IOSCO reports.
- OECD AI Principles.

Peer-reviewed financial, legal, and regulatory journal articles.

Publications of international financial organisations and regulatory agencies.
To capture recent advances in AI governance and financial regulation, only openly accessible and authoritative sources published primarily during 2020–2026 have been included.

3.3 Data collection and analysis

Systematic searches were carried out on recognised scientific online databases such as Scopus, Web of Science, SSRN, HeinOnline, Google Scholar, ScienceDirect and SpringerLink, as well as official publications from financial regulators and international organisations. The search terms included combinations of "artificial intelligence," "AI-generated forensic evidence," "market manipulation," "futures markets," "derivatives markets," "explainable AI," "algorithmic transparency," "evidentiary standards," "market surveillance," and "procedural fairness." The inclusion criteria were peer-reviewed academic studies, regulatory publications, and official institutional sources published in English and directly addressing derivatives or futures markets, AI-enabled financial surveillance, or evidentiary and explainability standards. Opinion pieces, non-English publications, duplicate records, and studies focusing on unrelated financial markets were excluded. An initial search identified 126 documents, of which 42 met the inclusion criteria and were retained for detailed analysis.

Qualitative content analysis and comparative legal analysis were used for analysing the collected materials. Each document was then analysed to identify references to AI explainability, algorithmic transparency, auditability, model validation, procedural fairness, evidentiary disclosure, and digital forensic standards. The coding framework comprised five analytical categories: Algorithmic Transparency, Reliability and Validation, Auditability and Documentation, Disclosure and Explainability, and Procedural Fairness. Each document or relevant passage was assigned to one or more categories according to its substantive focus; for example, provisions concerning model testing and error rates were coded under Reliability and Validation, whereas requirements concerning record keeping, traceability, and model logs were coded under Auditability and Documentation. For instance, a CFTC guidance passage referring to model-version logging or documentation requirements would be coded under "Auditability and Documentation."

The themes that have been identified were then aligned to five analytical dimensions:

- Algorithmic Transparency
- The reliability and validation of the model.
- Analyse the issue of audit trail and documentation.
- Disclosure and Explainability Requirements

Procedural Fairness/Due Process

A comparative analysis was later performed between the regulation of the three regulators to highlight the similarities in their principles, differences in regulation and existing gaps in evidence. Results of this analysis were used to create the proposed Evidentiary Explainability Framework (EEF) that is discussed in the discussion section.

Limitations:

There are a number of caveats to be noted. Firstly, because it is purely secondary legal and regulatory information, the study does not involve empirical testing of AI surveillance systems or proprietary regulatory algorithms. Technical architecture, training data, and operational performance data for regulatory AI models are typically kept private because they are directly tied to market surveillance activities.

Secondly, financial regulatory governance of AI is a dynamic field. Authorities like the CFTC, ESMA, and SEBI are constantly issuing new guidance that adapts to technological innovations, so future policy changes could alter policy expectations.

Third, the study deals specifically with AI-generated forensic evidence that is relevant in the context of enforcing regulations against market manipulation in futures and derivatives markets. While the list of principles of evidence for securities regulation, anti-money laundering investigations, and financial supervision in general may be applicable and similar to those proposed in this research, it is not included in the former.

While these are some of the limitations, the study offers a thorough analysis of one of the greatest challenges emerging in financial regulation at this juncture—from a legal and policy perspective—and outlines a workable evidentiary structure that can assist with the effective enforcement of financial laws and procedures, as well as with achieving the procedural fairness required for the process.

4. Results and Discussion

4.1 Comparative Assessment of existing Regulatory frameworks:

The comparative analysis shows that key financial regulators are increasingly aware of the significance of AI in identifying advanced forms of manipulation. Traditional rule-based monitoring systems often miss detecting types of suspicious behaviour prevalent in the trading world, such as spoofing, layering, wash trading, insider trading, cross-market manipulation, and synthetic-information attacks, which are easily detected by AI-assisted surveillance systems.

But there's a long way to go on the acceptance of AI's forensic evidence in court.

The Commodity Futures Trading Commission (CFTC) has taken the most formalized stance, focusing on the implementation of AI systems that are explainable, on the governance of models and on audit trails in AI-assisted compliance systems. The CFTC is aware that AI outputs that impact enforcement actions should be accompanied by sufficient documentation, controls and explainability. In the current system, however, there are no clear requirements for acceptable model accuracy, acceptable error rates, a minimum acceptable level of explainability, or what needs to be disclosed as part of an enforcement action.

The European Securities and Markets Authority (ESMA) is taking a more governance-centric approach, in the wake of the European Union's AI Act and Digital Operational Resilience Act (DORA). Key principles are transparency and accountability, as well as the need for human oversight and risk management for high-risk AI systems in European regulation. However, ESMA currently offers little guidance on the admissibility of AI-generated forensic outputs in derivatives market abuse investigations.

Review of Futures Markets (RFM)
ISSN: 1933-7116
Volume 34 Issue 2 (2026)
Website: https://fra.st

The Securities and Exchange Board of India (SEBI) has been increasingly adopting artificial intelligence (AI) in its market surveillance infrastructure. Indian exchanges have been using AI systems to identify AT, which are playing a crucial role in enhancing AT regulatory oversight. The use of AI systems by Indian exchanges to identify AT is well established and is playing a vital role in regulating AT. The rules and principles on explainability, algorithmic disclosure, independent validation and procedural safeguards, however, are not as developed in formal law as in the United States and Europe.

The comparative findings are summarised below.

Table 1. Comparative Assessment of AI Evidentiary Standards

Evidentiary Parameter	CFTC	ESMA	SEBI
AI surveillance adoption	High	High	High
Explainability requirement	Moderate	Moderate	Emerging
Audit trail requirement	Yes	Partial	Limited
Independent model validation	Limited	Limited	Minimal
Error-rate disclosure	Not specified	Not specified	Not specified
Confidence score reporting	Limited	Limited	No
Human oversight	Required	Required	Recommended
Evidentiary disclosure standard	Not standardized	Not standardized	Not standardized

The analysis suggests that while there has been substantial advancement in AI governance, none of the regulators has developed a comprehensive body of evidence to regulate the admissibility of AI-generated forensic evidence under the law.

Table 2. Comparative Regulatory Assessment of AI Evidentiary Standards

Evidentiary Parameter	CFTC	ESMA	SEBI
AI used for market surveillance	✓	✓	✓
Explainability requirement	Moderate	Moderate	Emerging
Human oversight	Required	Required	Recommended
Audit trail requirement	Yes	Partial	Limited
Independent model validation	Limited	Limited	Minimal

Review of Futures Markets (RFM)
ISSN: 1933-7116
Volume 34 Issue 2 (2026)
Website: https://fra.st

Evidentiary Parameter	CFTC	ESMA	SEBI
Error-rate disclosure	Not specified	Not specified	Not specified
Confidence score reporting	Limited	Limited	No
Standardised evidentiary framework	No	No	No

Explain the key findings from the comparison—highlight that although all three regulators are using AI, none has yet established a complete evidentiary framework for AI-generated forensic evidence.

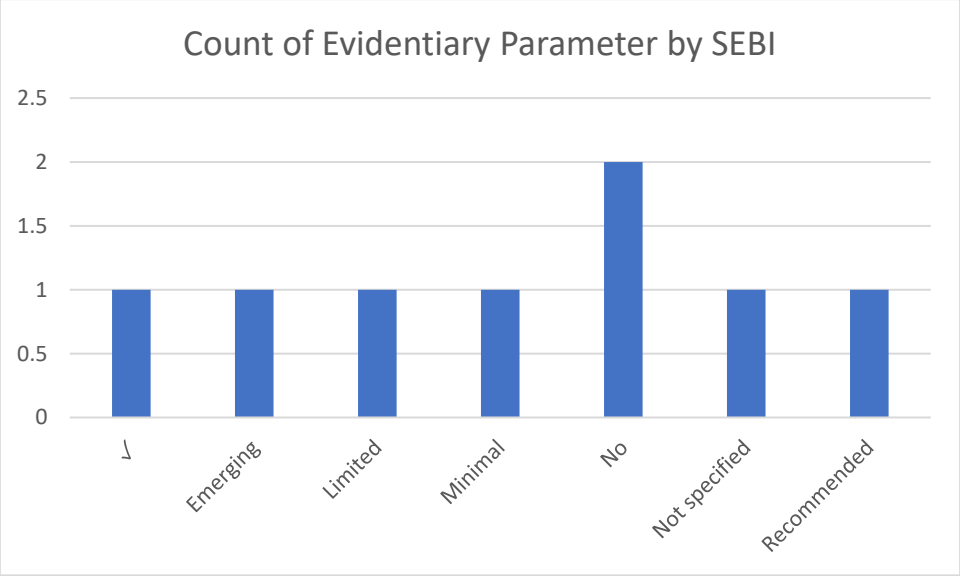

Figure 2. Distribution of Evidentiary Parameters under SEBI

The status of the various evidentiary parameters in SEBI's AI regulatory framework is depicted in Figure 2. Despite the uptake of AI surveillance, several important evidentiary principles have been either limited or not introduced, including disclosure of error rates and uniform requirements for AI evidence, which call for further regulatory guidance.

4.2 Major Evidentiary Challenges

The comparative analysis revealed similarities in problems that threaten the evidentiary reliability of AI-produced forensic evidence.

Black-Box Decision-Making

The majority of AI surveillance systems are based on deep learning and offer high accuracy but low explainability. Enforcement agencies can receive a high-confidence manipulation alert without a good understanding of the rationale behind the manipulation process. This will make it more difficult to be transparent and to facilitate judicial review.

The enforcement action in *United States v. Coscia* illustrates the importance of being able to reconstruct and explain the trading behaviour underlying a manipulation allegation. Although the case relied on conventional order-book and trading evidence rather than AI-generated

findings, it demonstrates the level of explanatory traceability that an AI-assisted spoofing determination would likewise need to provide.

An absence of standardised Explainability

The current regulations promote explainability but don't specify how much of it needs to be there. As a result, there may be varying levels of disclosure across different regulators in their evidence based on AI. This means that disclosures may vary across regulators.

The Coscia and Sarao spoofing matters further illustrate this concern because manipulation findings could be linked to identifiable order-placement, cancellation, and trading patterns. If comparable findings are generated by machine-learning surveillance, regulators would need sufficiently standardised explanations to show how the system connected the observed behaviour to a manipulation alert.

Model Validation

Most regulatory guidance calls for model testing but does not provide any minimal validation procedures. While AI-generated evidence can be valuable in a court case, several questions remain to be addressed before it can be relied upon. There are a number of questions that need to be answered before AI-generated evidence can be used as forensic evidence in a court case, including questions about dataset quality, overfitting, model drift, and external validation.

The Sarao enforcement matter demonstrates why validation becomes particularly important when surveillance must distinguish manipulative layering or spoofing from legitimate high-frequency order activity. An AI model used for a comparable investigation would therefore require documented testing to demonstrate that its classification is sufficiently reliable and is not merely capturing unusual but lawful trading behaviour.

Error rates of no items were checked

In most cases, the traditional forensic sciences will reveal their known error rates when making expert testimony. Adjudicators have a hard time determining the level of reliability when the forensic output is produced by AI since such output rarely contains information about false positives, false negatives, confidence intervals, and uncertainty estimates. In a spoofing investigation comparable to Coscia, an AI-generated manipulation flag could have significant evidentiary consequences. Disclosure of false-positive and false-negative rates would therefore assist adjudicators in determining whether the algorithm reliably distinguishes intentional spoofing from legitimate order cancellation and modification.

Limited Auditability

Forensic evidence should be well documented, showing the work which led to the conclusion. Therefore, AI systems should maintain detailed logs of model versions, input data, preprocessing methods, model parameters and generated outputs. There is currently no uniformity in regulatory requirements in this regard. The detailed reconstruction of order placement and cancellation activity in the Coscia and Sarao matters illustrates the importance of a traceable evidentiary record. Where AI performs part of this analytical reconstruction, equivalent auditability would require preservation of relevant inputs, model versions, processing steps, outputs, and subsequent human review.

Review of Futures Markets (RFM)
ISSN: 1933-7116
Volume 34 Issue 2 (2026)
Website: https://fra.st

Procedural Fairness

Perhaps the most important question is a due process one. At the very least, there should be meaningful opportunities for those who are subject to enforcement action to review and contest algorithmic evidence. An absence of transparency means that the other party to the administrative procedure can't effectively challenge the assumptions of the model, training data or the methodology used in analyses, which can impact the fairness of the administrative procedure. The Coscia and Sarao cases also demonstrate the importance of allowing allegations of manipulative trading to be tested against an identifiable evidentiary record. If an enforcement authority were instead to rely materially on an AI-generated classification, procedural fairness would require the affected party to receive sufficient information to understand and meaningfully challenge the basis, reliability, and limitations of that algorithmic finding.

4.3 Evidentiary Explainability Framework (EEF)

The comparative analysis highlights the need to recognise the significance of explainable AI, but lacks a clear and detailed set of guidelines for the evidence produced by AI in forensics. To fill this void, the current study aims to suggest an Evidentiary Explainability Framework particularly designed for derivatives marketplace manipulation enforcement. The framework aims to mitigate potential issues with technological innovation and accountability through transparency, reliability, and procedural fairness of AI-powered enforcement.

The EEF has five interrelated pillars focusing on key aspects of evidence reliability.

Pillar I: Algorithmic Transparency

The first pillar calls for regulators to make adequate information public about the AI model that is used for market surveillance. While proprietary algorithms do not have to be fully disclosed, meaningful explanations of the model's purpose, analytical methodology, key variables, decision criteria, and limitations should be provided to enable enforcement agencies to offer a meaningful explanation of the model. This openness helps ensure that conclusions drawn by algorithms are understandable to adjudicators and those affected.

If the AI surveillance system is used to detect spoofing, the regulator should clarify how it drew its conclusions, such as if an order was cancelled repeatedly, the order book showed unusual dynamics, the execution time was unusual, there were unusual trading patterns, or the behaviour was similar to an order in a case that was previously classified as manipulation.

The proposed EEF addresses the principal shortcomings identified in existing regulatory frameworks by introducing five interconnected disclosure requirements. Together, these pillars provide a structured basis for evaluating the legal reliability and admissibility of AI-generated forensic evidence in derivatives market enforcement.

Comparison of the EEF with Existing AI Governance Frameworks

Review of Futures Markets (RFM)
ISSN: 1933-7116
Volume 34 Issue 2 (2026)
Website: https://fra.st

EEF Pillar	Corresponding NIST AI RMF Function	Relevant EU AI Act Provision	Relationship
Algorithmic Transparency	Govern / Map	Articles 13–14	Supports transparency, information provision, and meaningful human oversight of AI systems.
Reliability and Error Disclosure	Measure / Manage	Article 15	Aligns with requirements concerning accuracy, robustness, cybersecurity, and performance monitoring.
Auditability and Documentation	Govern / Measure	Articles 11–12	Supports technical documentation, record keeping, traceability, and review of AI-system operation.
Independent Validation	Measure / Manage	Articles 9 and 15	Relates to risk management, testing, validation, accuracy, robustness, and continuing assessment.
Procedural Fairness	Govern / Manage	Articles 13–14	Reinforces transparency and human oversight, while the EEF extends these concepts toward the ability to understand and challenge AI-generated forensic findings.

The comparison demonstrates that the proposed EEF is broadly compatible with established AI governance principles while addressing a distinct evidentiary objective. The NIST AI RMF functions of Govern, Map, Measure, and Manage provide a general risk-management structure, while Articles 9–15 of the EU AI Act establish requirements concerning risk management, data governance, technical documentation, record keeping, transparency, human oversight, accuracy, robustness, and cybersecurity. The EEF builds on these governance principles by translating them into requirements relevant to the reliability, disclosure, auditability, validation, and contestability of AI-generated forensic evidence in derivatives-market enforcement.

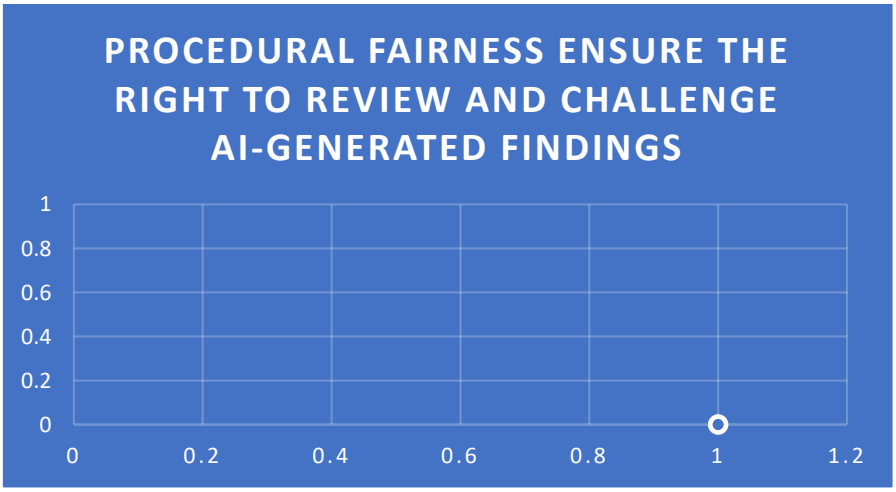

Figure 3. Procedural Fairness in AI-Generated Forensic Evidence,

Figure 3 highlights the importance of procedural fairness in AI-assisted market enforcement. It emphasises that individuals should have the right to review, understand, and challenge AI-generated findings, ensuring transparency, accountability, and due process in regulatory decision-making. Source: Author's conceptualisation, not derived from empirical data**.**

Pillar II: Disclosure of errors and reliability

One of the basic expectations of any type of forensic evidence is its reliability. AI-generated outputs should therefore be accompanied by objective measures of the model's performance to show its predictive ability.

These should be measures such as:

- Model accuracy
- Precision
- Recall
- False positive rate
- False negative rate
- Confidence score
- Validation results

Making these metrics available gives adjudicators a way to determine the confidence in AI-generated conclusions without taking them on faith, as they are algorithmic results. This requirement for disclosure is already in place in a number of forensic science disciplines, and should be broadened to cover AI-assisted financial enforcement.

Pillar III: Auditing and reproducibility

All evidence generated by AI should be 100% reproducible. All major processing steps of the algorithms need to be recorded, with detailed audit trails.

There should be a suitable audit trail of:

- Input datasets
- Data preprocessing procedures
- Model version
- Parameter configuration
- Time of execution
- Software environment
- Generated outputs
- Human review undertaken

The comprehensive documentation means that it is possible for the analytical process to be repeated by an independent expert if necessary in an administrative and/or legal review.

Pillar IV: Independent Validation

To ensure the credibility of AI-generated forensic evidence, it is crucial for it to be verified independently. Surveillance algorithms should be periodically subjected to independent technical audits using suitable metrics and competent external experts by regulatory agencies.

Validation should assess:

- Robustness
- Generalizability

Review of Futures Markets (RFM)
ISSN: 1933-7116
Volume 34 Issue 2 (2026)
Website: https://fra.st

- Dataset quality
- Algorithmic bias
- Stability over time
- Cybersecurity resilience

Independent certification helps mitigate the risk of institutional bias, thereby building public trust in AI-aided enforcement.

Pillar V: Procedural Fairness

The last pillar is that AI should complement, not supplant, human decision-making.

Those who are subject to enforcement action should be given the chance to:

- Review AI-generated findings.
- Examine supporting documentation.
- Challenge algorithmic assumptions.
- Question model reliability.
- Produce unbiased expert testimony.

Ask for clarification about Explainability Reports.

Keeping human oversight in meaningful control of AI helps ensure basic principles of administrative justice, and that AI doesn't, and should never, make decisions for itself.

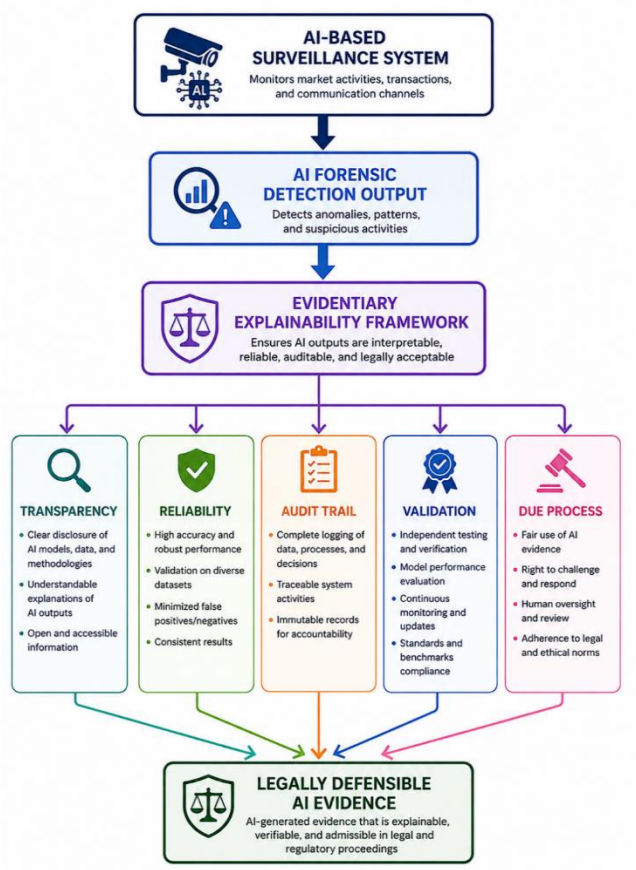

Figure 4. Proposed Evidentiary Explainability Framework (EEF), Source: Drawn by Author

Feature Importance Error Rate Reproducibility Bias Testing Human Oversight

By outlining a structure for regulators to consider when assessing whether AI-generated forensic products meet minimum evidentiary standards, the framework offers a useful tool for determining whether AI can be used in the future to assist in market-manipulation investigations.

4.4 Policy Implications

This study has significant implications for both financial regulators and adjudicators, financial institutions, and AI developers.

The proposed framework provides a harmonised approach for assessing forensic evidence generated by AI tools, which is beneficial for regulators like the CFTC, ESMA and SEBI. The regulatory bodies should set measurable disclosure requirements (model validation, reporting uncertainty, audit trails, independent verification and certification, etc.) instead of just explainability requirements.

Administrative tribunals and courts could also have more defined guidelines on the evidence in AI-enforcement cases. Consistent and uniform disclosure of algorithmic information would help to streamline judicial decision-making and minimise the risk of uncertainty about the admissibility of algorithmic evidence. AI-based compliance solutions should include explainability, documentation, and auditability from the beginning of model development. By fostering responsible AI governance, compliance is enhanced, and organisational risk management and stakeholder trust are boosted.

5. Conclusion

Artificial intelligence (AI) has become a part and parcel of modern-day derivatives market surveillance. New machine learning algorithms can highlight manipulation methodologies that are too complex to be identified by conventional surveillance systems. While AI can serve as a valuable investigative tool, it is also becoming increasingly important that outputs from artificial intelligence are regarded as evidence in investigations today. Currently, as AI reshapes financial regulation, AI outputs are being treated as forensic evidence, not just investigative tools. This change raises some key legal issues relating to transparency, explainability, reliability and procedural fairness.

The present study focused on the evidentiary implications of a comparative analysis of regulatory developments in the U.S., the European Union, and India regarding the enforcement of market manipulation in futures and derivatives markets. The analysis highlights that the importance of "explainable" AI is gaining visibility at the CFTC, ESMA, and SEBI, but current regulations remain lacking. Presently, there is no uniform guidance on which aspects of algorithms must be disclosed, validated, reported as uncertainties, independently verified, or protected by procedural safeguards. There is currently no uniform guidance on what should be disclosed, validated, reported as uncertainties, independently verified, or protected by procedural safeguards regarding aspects of algorithms.

The results show that explainability does not necessarily lead to legality of the forensic evidence produced by AI. Improved evidentiary governance needs a larger ecosystem encompassing the following: (1) Transparency of algorithms, (2) objective performance evaluation, (3) full documentation, (4) independent technical validation and (5) meaningful

opportunities for challengers to question the algorithmic result. If these fail to be in place, over-dependence on opaque AI systems could jeopardise the enforceability of the financial markets, their regulatory legitimacy, and public trust in the process.

This paper introduced the Evidentiary Explainability Framework (EEF), which comprises five interrelated pillars: algorithmic transparency, reliability and error disclosure, audit trail preservation, independent validation, and procedural fairness. The framework offers a structured approach for regulators to consider when assessing the minimum evidentiary standards of AI-generated forensic outputs to be used during an enforcement action. The framework builds on the principles of explainable AI, traditional digital evidence and administrative law, and enhances the trustworthiness and accountability of AI-driven regulation.

The more financial markets evolve towards increased automation, the more the harmonisation of AI evidentiary standards from around the world will become relevant. Regulatory action in the future should go beyond the basic tenets of responsible AI to tangible disclosure requirements that can help ensure that technological innovation is not at odds with due process, transparency, and the rule of law. Such changes will allow regulators to benefit from the analytical power of AI and maintain the fairness and legitimacy that is so critical to good financial market governance.

References:

1. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052
2. Aggarwal, R., & Wu, G. (2006). Stock market manipulations. *The Journal of Business, 79*(4), 1915–1953. https://doi.org/10.1086/503655
3. Comerton-Forde, C., & Putniņš, T. J. (2011). Measuring closing price manipulation. *Journal of Financial Intermediation, 20*(2), 135–158. https://doi.org/10.1016/j.jfi.2010.04.001
4. Doshi-Velez, F., & Kim, B. (2017). *Towards a rigorous science of interpretable machine learning*. arXiv. https://arxiv.org/abs/1702.08608
5. European Parliament and Council of the European Union. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj
6. European Securities and Markets Authority. (2024). *Statement on the use of artificial intelligence in investment services*. https://www.esma.europa.eu
7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://www.deeplearningbook.org
8. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys, 51*(5), Article 93. https://doi.org/10.1145/3236009

9. International Organization of Securities Commissions. (2021). *Artificial intelligence and machine learning in capital markets: Use cases and challenges.* https://www.iosco.org
10. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*. https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
11. Organisation for Economic Co-operation and Development. (2019). *OECD principles on artificial intelligence.* https://oecd.ai/en/ai-principles
12. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778
13. 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. https://doi.org/10.1038/s42256-019-0048-x
14. United States Commodity Futures Trading Commission. (2024). *Technology Advisory Committee advances report and recommendations to the CFTC on responsible artificial intelligence in financial markets.* https://www.cftc.gov/PressRoom/PressReleases/8905-24
15. United States Commodity Futures Trading Commission. (2025). *Criminals increasing use of generative AI to commit fraud.* https://www.cftc.gov/LearnAndProtect/AdvisoriesAndArticles/AI_Fraud.html
16. United States National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0).* https://doi.org/10.6028/NIST.AI.100-1
17. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012
18. Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 80–89. https://doi.org/10.1109/DSAA.2018.00018