Background

In light of increased adoption of artificial intelligence (AI) and machine learning (ML) technologies by market participants, on 20 June 2025, the Securities and Exchange Board of India (SEBI) issued a consultation paper titled "*Guidelines for Responsible Usage of AI/ML in Indian Securities Markets*" (Paper). The Paper recognises that while AI/ML holds immense potential to augment market efficiency, facilitate complex decision-making, and bolster regulatory investigations through the analysis of large datasets, it also gives rise to certain risks. Given the scale, pace, and impact of AI/ML-driven decisions in financial markets, any misuse or malfunction could have far-reaching consequences for market integrity and investor protection. In this context, the Paper seeks to establish guiding principles that reconcile innovation using AI/ML with safeguards for investor interests to preserve the integrity of the securities market. Stakeholders can submit their comments on the Paper until 11 July 2025.

Key Recommendations

At the outset, the Paper acknowledges that market participants are involving AI/ML applications in a plethora of use cases, *inter alia*, including advisory and support services, risk management, and client identification. In this context, the Paper considers global best practices to outline high-level principles that should underpin the governance of AI/ML applications in the securities market:

(i) **Model Governance**: Market participants using AI/ML are expected to have skilled internal teams for effective human oversight over AI/ML deployments. They are also expected to ensure robust governance, fallback plans, and execute robust agreements while engaging third-party vendors/service providers. Continuous monitoring, independent audits, and periodic reporting of accuracy results of AI/ML models to SEBI are some of the other requirements envisaged in the Paper.

(ii) **Investor Protection and Disclosures**: Market participants using AI/ML models in business operations that directly impact customers (for instance, algorithmic trading or advisory services) are also expected to make certain disclosures to clients to ensure trust, transparency, and accountability. Disclosures are expected to be in simple language and are required to cover, *inter alia*, product features, purpose, risks, model accuracy, and fees. Additionally, investor grievance mechanisms for AI/ML systems are expected to comply with SEBI's existing regulatory framework in this regard.

(iii) **Testing Framework**: Market participants should test AI/ML models in a segregated environment prior to deployment. This is to ensure that AI/ML models behave as expected in both stressed as well as unstressed market conditions. Market participants should also maintain proper documentation of all the models and store input and output data for at least five years. Additionally, market participants are also expected to document the logic of AI/ML models to ensure that the outcomes produced are explainable, traceable and repeatable. Notably, in addition to the existing methods of testing, market participants are expected to perform shadow testing with live traffic of AI/ML models to ensure quality and performance before deployment in the production environment. This indicates a notable shift from one-time pre-deployment testing towards a lifecycle-oriented, real-time validation approach, to keep pace with evolving model behaviour in dynamic market conditions.

(iv) **Fairness and Bias**: Market participants should implement appropriate processes and controls to identify and remove biases from data sets (i.e., not favour or discriminate one group of clients/customers over another). From a business standpoint, while the Paper remains silent on whether objective, reasonable classification between distinct customer groups is envisaged as permitted, such differentiation, when based on legitimate and non-discriminatory criteria, may arguably be conceived as allowed. Without a clear definition of what constitutes 'fairness', businesses may practically need to conduct fairness impact assessments towards auditing the fairness of the outcomes of their AI/ML applications as part of their overarching AI governance framework.

(v) **Data Privacy and Cybersecurity**: Market participants using AI/ML systems are required to establish clear policies for data security and cybersecurity, and are required to ensure that the collection, use, and processing of personal data of investors adheres with applicable data protection laws. Market participants are also expected to promptly report any technical glitches or data breaches to SEBI and other relevant authorities. This signals toward increased regulatory convergence between sector-specific regulation and broader requirements under data protection and cybersecurity laws, paving the way for future coordination between the SEBI, and authorities such as the Indian Computer Emergency Response Team, and the upcoming Data Protection Board of India, particularly in critical aspects such as the reporting of cyber security incidents and personal data breaches.

Risks and Control Measures

The Paper additionally annexes a checklist for managing anticipated threats posed in the context of AI/ML applications, identifying six key categories of risk, and corresponding mitigation strategies:

(i) **Malicious Use**: The Paper identifies the capability of Generative AI to fabricate fraudulent financial statements, misleading news articles, or deepfake content, which may potentially lead to price manipulation or market instability. To address this concern, it recommends (a) watermarking and provenance tracking; (b) reporting of suspicious activities by market participants; and (c) educating investors about AI-generated misinformation risks through public awareness campaigns.

(ii) **Concentration Risks**: The Paper highlights that relying on a limited number of Generative AI providers by market participants could contribute to systemic risks in times of failure or impairment. To this end, the Paper suggests (a) proactive monitoring of market concentration; (b) diversification of service providers; and (c) enhanced monitoring of critical vendors and their AI applications/tools.

(iii) **Herding and Collusive Behaviour**: The Paper flags the risk associated with herding and collusive behaviour attributable to the overlapping use of similar AI models or datasets, especially by large or systemically important market participants. To mitigate herding or collusive behaviour arising from converging AI strategies, the Paper espouses (a) diversity in AI architectures and data sources; (b) monitoring stock exchanges to identify potential herding behaviour; (c) conducting regular algorithmic audits to detect collusive patterns; and (d) deploying circuit breakers to respond to market volatility amplified or driven by AI/ML applications.

(iv) **Lack of Explainability**: The Paper acknowledges the challenge of explainability in AI systems and suggests measures to ensure transparency and meaningful oversight. It recommends (a) mandating market participants to document AI processes in detail; (b) encouraging the use of interpretable AI models or explainability tools which may aid in deciphering the logic or working of an AI/ML model; and (c) mandating human review of AI-generated outputs.

(v) **Model Failure / Runaway Behaviour**: The Paper acknowledges that flaws in AI/ML applications could cause financial instability. In light of this, the Paper recommends (a) stress testing to assess the performance of AI systems in extreme scenarios; (b) volatility controls through kill switches and circuit breakers; and (c) human oversight to control the over-reliance on AI systems and establish clear lines of human accountability for AI-driven decisions.

(vi) **Lack of Accountability and Regulatory Non-Compliance**: The Paper also highlights the risk of regulatory infractions and investor losses stemming from the unaccountable use of AI systems, particularly, where the outcome of these systems is bereft of effective monitoring. Notably, the Paper also highlights the risk of market participants attempting to avoid liability for AI-driven outcomes by attributing them to AI systems. To address this, it recommends (a) testing AI tools in regulatory sandboxes; (b) training staff to understand and manage compliance risks linked to AI deployment; as well as (c) human-in-the-loop or human-around-the-loop mechanisms.

Comments

Although India is yet to adopt an overarching framework for AI governance, the Paper demonstrates SEBI's emerging role as an early mover in shaping practical guardrails for AI/ML use in financial markets. By outlining expectations around the testing of AI/ML applications, as well as steps to ensure fairness, and human accountability, SEBI is laying the groundwork for responsible AI adoption in financial markets, paving the way for an overarching, cross-sectoral legal framework.

- *Supratim Chakraborty (Partner); Tomu Francis (Partner); Siddharth Sonkar (Senior Associate) and Mayank Barman (Associate)*

About Khaitan & Co

Khaitan & Co is a top tier and full-service law firm with over 1200 legal professionals, including 300+ leaders and presence in India and Singapore. With more than a century of experience in practicing law, we offer end-to-end legal solutions in diverse practice areas to our clients across the world. We have a team of highly motivated and dynamic professionals delivering outstanding client service and expert legal advice across a wide gamut of sectors and industries.

To know more, visit **www.khaitanco.com**