Brussels, 11 April 2025

EFAMA Feedback on IOSCO's Report: Artificial Intelligence in Capital Markets

Background information

This comment letter is intended to be a public comment on IOSCO's Report Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges

EFAMA leads its work on artificial intelligence through its Artificial Intelligence Taskforce (AI TF), which supports asset managers in navigating both the opportunities and challenges posed by AI technologies. The AI TF also plays a key role in guiding EFAMA members in implementing the EU AI Act.

EFAMA published its AI System Assessment Tool earlier this year, which is publicly available on our website. The tool is designed to help firms of all sizes document and classify their AI applications in line with the EU AI Act's risk-based framework. It can be accessed at the following link: EFAMA AI System Assessment Tool.

We welcome the publication of IOSCO's *Report on AI in Capital Markets* ('the Report'). EFAMA's AI TF has reviewed the report in detail and launched an internal survey to gather member feedback. The views presented in this response reflect the outcomes of that survey and the discussions that followed within our group.

Risk Areas Identified by the Report

Among the five risk areas identified by IOSCO in its report — namely (i) malicious uses, (ii) risks related to AI models, (iii) data considerations, (iv) concentration, outsourcing, and third-party dependency, and (v) interactions between humans and AI — our members identified interactions between humans and AI as the highest risk related to the use of AI in financial markets. In particular, respondents recognised insufficient oversight and over-reliance on AI for decision-making as the most pressing issues.

This was followed by malicious uses of AI (i.e. cybersecurity threats and fraud), which were generally rated as medium to high risk. Risks associated with AI models and data considerations were perceived as moderate — present but manageable — while concentration, outsourcing, and third-party dependency were consistently rated as the lowest areas of concern among respondents.

A key observation from these findings is that our group — composed predominantly of practitioners closely involved in the operational implementation of the EU AI Act within their firms — tends to prioritise practical and operational risks over more structural or theoretical ones. In this respect, the group appears to have adopted a 'probabilistic' approach in rating these risks, giving greater importance to the likelihood of a risk materialising when determining where risk management and regulatory focus should be directed. When

asked to comment on the section of the Report concerning *Market Dynamics, Potential Outcomes, and Data and Knowledge Gaps* — which includes areas such as interdependence, herding effects, and 'collusive' or 'scheming' behaviours — most members did not view these as priority concerns and generally ranked them lower in terms of perceived risk. While none of these areas were dismissed entirely, they were not prioritised as key concerns when compared to more operational risks.

The strong concern around human-AI interaction should serve as a basis for firms to move beyond technical compliance and address this issue by reviewing their organisational processes and providing meaningful and relevant staff training. Such training should ideally be aligned with the level and role that an employee has in relation to the AI system that is being used. A concept that frequently comes up during our discussions is the distinction between having a '*human in the loop*' and an '*expert in the loop*'. While both imply a level of human oversight of AI outputs, the latter refers to an individual with the relevant expertise to critically assess and effectively apply those outputs.

Although this distinction is essential, it should not lead to the assumption that expertise alone guarantees robust oversight. Automation bias and cognitive distortions can affect even highly trained professionals. Experts may overestimate their ability to interpret or challenge AI-generated results—especially when model behaviour is opaque. Furthermore, repeated exposure to seemingly accurate AI outputs can lead to complacency or growing over-reliance over time. This dimension of behavioural risk should be taken into consideration when approaching the deployment of AI in financial services (see our reference below to the 'black box' effect and the importance of interpretability and explainability of AI systems)

We therefore encourage firms to adopt proactive measures to address these risks. These should include not only regular training (across different departments) but also the implementation of internal controls and oversight. For example, some firms have employed 'false positives' (i.e. deliberately inserting inaccurate or misleading AI-generated suggestions) to test whether their staff is able to detect and challenge faulty outputs. Such safeguards are crucial in mitigating automation bias, promoting critical engagement, and fostering a culture of responsible AI use.

As previously mentioned, *malicious uses* were identified by EFAMA's AI TF as the second highest-ranking risk area related to AI in financial services (rated on average as medium to high risk). Malicious actors are expected to increasingly leverage AI tools to enhance the effectiveness and scale of harmful or criminal activities. Asset management firms — and financial services more broadly — are exposed to these evolving threats and will need to implement appropriate safeguards to mitigate the associated risks.

Examples of AI-enabled cyber threats include the use of malicious large language models (LLMs) to facilitate phishing attacks, the spread of misinformation that could damage the reputation of firms or individuals, voice cloning to bypass personal identification procedures, deepfakes to impersonate intermediaries, and broader information security concerns — such as the inadvertent storage of confidential data by publicly accessible LLMs, which may be reused in future outputs[1].

Given these risks, we recommend that firms review and update their risk management frameworks to account for this new dimension of malicious activity. Furthermore, we encourage supervisory authorities to allocate sufficient resources and attention to this risk area, particularly when considering further initiatives in the development of tools and guidance for regulators and firms, investor and user education, and enhanced supervisory cooperation.

Moreover, we consider that the *interpretability and explainability* of AI systems—falling under the category of risks related to *AI Models* within the Report's proposed taxonomy—should not be underestimated as a risk factor. When our members were surveyed on the so-called *'black box'* effect of AI systems (i.e. the complexity and opacity of AI, which makes it difficult to understand how outputs are generated), the majority agreed that this poses a risk worth considering. Our members consider that it becomes more difficult to

[1] The Investment Association (The IA), 'AI Cyber Risk: Immediate Actions for Firms – Recommendations from the IA Cyber Resilience Committee'.

diagnose errors without adequate transparency, which can weaken trust in oversight and governance mechanisms.

While opinions varied on the degree of concern—some viewed it as a *moderate risk* that can be managed through proper oversight and safeguards, while others regarded it as a *significant risk* with serious implications for transparency and accountability—there was broad consensus that it should be continuously monitored. The relevance of this issue is likely to increase as more advanced AI systems (i.e. deep learning models and LLMs) are deployed in financial markets. Our recommendations on how IOSCO may address the risks surrounding interpretability and explainability of AI systems—through the promotion of supervisory principles—are outlined in the following section.

Finally, we suggest that IOSCO consider addressing *Autonomous AI Agents*—systems capable of making decisions, initiating actions, and interacting with environments independently—as a distinct and emerging risk category within its methodology. Unlike traditional models that support human users, these agents operate with goal-oriented autonomy, which:

- increases the risk of unintended behaviours and runaway dynamics;
- introduces new forms of agent-to-agent interaction, potentially amplifying market effects; and
- blurs accountability structures, particularly when outputs are neither easily traceable nor explainable.

Given their growing role in financial markets—such as trading, compliance workflows, or client interaction—these systems may warrant separate risk classification and governance expectations beyond what is currently covered under general model or outsourcing risks. The following section outlines our recommendations on how IOSCO may address the issue of Autonomous AI Agents through the promotion of supervisory principles.

Recommendations for Future IOSCO Work in AI

We welcome IOSCO's initiative to outline possible directions for the second phase of its work on artificial intelligence. To inform our response, we surveyed our members on the potential future areas of work identified by IOSCO in its report, namely: (i) the development of additional tools, good practices, or guidance; (ii) investor education; (iii) cross-border information sharing on key risks; (iv) supervisory cooperation; and (v) capacity building and technical assistance in support of IOSCO members.

Our members were asked to indicate which potential IOSCO workstreams should be prioritised. The highest-ranking priorities were:

a. the development of tools, recommendations, or principles to help IOSCO members address the challenges, risks, and issues posed by the use of AI in financial products and services; and

b. investor education, particularly regarding the growing number of frauds involving the use of AI.

There was also strong support for IOSCO to work towards developing *international standards and minimum principles* for the deployment of AI in the financial sector. These should be guided by *technological neutrality* and should aim to provide *legal certainty*, particularly in relation to how sector-specific obligations apply when AI tools are used to deliver financial products and services.

In this respect, we would like to emphasise that for asset management firms the use of AI is *ancillary* to their primary business — which is the management of funds and portfolios. While this may seem self-evident, it is an essential consideration that should underpin any future regulatory approach to AI in financial services. AI should be recognised as a supportive technology rather than the core business activity of our member firms. As such, regulatory frameworks should ensure that the compliance burden is proportionate and does not discourage innovation or the deployment of AI tools. This point is likely relevant not only to asset management but also to other segments of the financial services industry.

Some of the issues that IOSCO could address through the development of international standards and minimum principles include:

1. Autonomous AI Agents. Given the potential for Autonomous AI Agents to have an increasingly present role in trading, compliance, and customer service, we recommend IOSCO to develop supervisory principles specifically for AI agents that operate with goal-oriented autonomy. These principles could include thresholds for intervention, escalation protocols, and rules for managing agent-to-agent interactions.

2. Interpretability and Explainability of AI Systems. IOSCO could take the lead in establishing global baseline expectations for model transparency, particularly in relation to high-impact AI use cases in financial services. These may include:

- A risk-tiered framework for explainability requirements (i.e. applicable standards for suitability assessments, fraud detection, or automated decision-making);
- Guidance distinguishing between intrinsic model transparency and post hoc explainability tools;
- Clarification on how explainability relates to regulatory auditability and accountability—particularly under overlapping frameworks such as MiFID II, DORA, and the EU AI Act.

Within the European asset management industry, significant uncertainty remains regarding how the provisions of the EU AI Act will interact with existing financial regulations, including MiFID II, GDPR, and DORA. Firms subject to these regimes are already required to comply with extensive governance and risk-management rules. Many of the obligations under the EU AI Act therefore appear to overlap with — or duplicate — existing requirements. We anticipate that similar issues will arise in other jurisdictions as regulatory frameworks for AI are developed and enacted. In light of this, we encourage IOSCO to consider this dimension when developing international recommendations and standards. In particular, future guidance should ensure that AI-specific frameworks are streamlined and integrated into existing financial regulation, thereby simplifying compliance and promoting innovation. This is especially important in the context of horizontal legislation, such as the EU AI Act, which applies across all sectors — including many with vastly different levels of AI maturity and risk.

Furthermore, there are several additional areas of legal uncertainty surrounding the use of AI that IOSCO could address through further work on tools, good practices, or guidance. These include, for example, liability, intellectual property, copyright and transparency requirements — particularly in the context of generative AI models. For instance, ESMA issued guidance to firms using artificial intelligence in investment services (May 2024), focusing on the use of AI by investment firms and its interaction with relevant MiFID II requirements[2], such as the imperative to always prioritise their client's best interests. We welcome guidance issued by regulatory bodies in this area, as it helps firms in the financial services industry more easily navigate the complexities of AI within the broader framework of financial regulation to which they are subject.

Regarding investor education, we strongly encourage IOSCO to further its work in this area. As previously noted, EFAMA members identified human interaction with AI and malicious use cases as the two risk areas with the highest relevance. Investor education initiatives could play a vital role in not only preventing scams but also in raising awareness about how AI is being used in financial services, ultimately enhancing investor trust in financial markets. For this reason, we recommend that IOSCO prioritise its educational initiatives in these two areas.

Additionally, we suggest that IOSCO develops educational resources targeted at both investors and financial services firms — i.e. users of AI tools. Both audiences would benefit from a clearer understanding of how to manage risks associated with human-AI interaction and the malicious use of AI. We believe these efforts could meaningfully support the safe and responsible adoption of AI technologies in financial services.

As a final point, our members identified supervisory cooperation and cross-border information sharing as areas of potential value for IOSCO to further its work, particularly in relation to addressing the risks posed by malicious uses of AI. It is important to highlight that mitigating such risks could benefit significantly from enhanced international collaboration. In particular, IOSCO could play a leading role in strengthening

[2] European Securities and Markets Authority (ESMA), 'Public Statement on the Use of Artificial Intelligence in the Provision of Retail Investment Services' (May 2024).

cooperation across jurisdictions by facilitating the sharing of insights, identifying emerging threats, and coordinating responses to AI-related risks. Supervisory cooperation will also be crucial in enhancing cross-border supervision and enforcement, particularly in instances where AI is utilised by firms and service providers operating across multiple regulatory frameworks.

Given that many AI-enabled frauds and scams are likely to be transnational in nature, it is essential that governments and supervisory authorities coordinate their efforts. In this context, we believe it is equally important to establish a structured dialogue with the industry to understand the practical challenges better and to ensure that supervisory approaches are both effective and proportionate to the nature of the threats. Such collaboration would greatly support the development of a coherent and responsive international framework for addressing malicious AI use targeted at financial markets. Especially firms operating in a cross-border context require a consistent and at least non-contradictory regulatory framework to drive innovation further.

ABOUT EFAMA

EFAMA is the voice of the European investment management industry, which manages EUR 28.5 trillion of assets on behalf of its clients in Europe and around the world. We advocate for a regulatory environment that supports our industry's crucial role in steering capital towards investments for a sustainable future and providing long-term value for investors. Besides fostering a Capital Markets Union, consumer empowerment and sustainable finance in Europe, we also support open and well-functioning global capital markets and engage with international standard setters and relevant third-country authorities. EFAMA is a primary source of industry statistical data and issues regular publications, including Market Insights and the EFAMA Fact Book. More information is available at www.efama.org

Contact:
Franco Luciano
Regulatory Policy Advisor, Capital Markets & Digital
Franco.luciano@efama.org | +32 4 902 774 79