Supervisory Toolkit for AI Use in Capital Markets

FINAL REPORT

The Board of the
International Organization of Securities Commissions

FR/02/2026 May 2026

Copies of publications are available from
The International Organization of Securities Commissions website

iosco.org

© International Organization of Securities Commissions 2026.
All rights reserved. Brief excerpts may be reproduced or translated
provided the source is stated.

Table of Contents

Executive Summary

This report supports IOSCO member authorities in their respective oversight of the use of Artificial Intelligence (AI) systems by entities subject to their regulation and supervision[1] through a toolkit that provides supervisors with practical, non-binding, non-prescriptive supervisory tools, applicable across regulatory models. This report is the result of a multi-phased approach by IOSCO through its Fintech Task Force (FTF) to assist IOSCO members as they consider regulatory and supervisory responses to AI technologies used in capital markets. This approach is based on a shared understanding of the risks such technologies may pose to investor protection, market integrity, and financial stability. The next phase of IOSCO's AI work will be to conduct a review of emerging industry practices across the areas of supervisory focus noted in the report. IOSCO will continue to play a coordinating role addressing AI developments in capital markets and will continue to engage with other relevant international organizations such as the Financial Stability Board (FSB).

IOSCO's existing work on AI includes "The use of artificial intelligence and machine learning by market intermediaries and asset managers"[2] (the 2021 AI Report), and "Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges"[3] (the 2025 AI Report). Throughout IOSCO's work, AI systems have continued to undergo significant developments, including through the incorporation of Generative AI (GenAI), and the emergence of Agentic AI techniques. Ongoing advancements in AI technologies are expanding the range of AI applications in capital markets, bringing potentially transformative benefits but also introducing or amplifying potential risks.

These risks, outlined in the 2025 AI Report, underscore the need for adaptive supervisory approaches, especially given the increasing complexity and opacity of certain AI systems, the continuing evolution of AI use cases, and the potential for concentrations in and dependencies on AI service providers. The toolkit in this report aims to cover the full lifecycle of a particular AI system and to apply to all AI system types, from those based on more traditional forms of Machine Learning (ML) to those using GenAI and emerging Agentic AI techniques. The report draws on responses to a recent IOSCO FTF Member survey, prior IOSCO work,

1 Referred to as supervised firms, or firms, throughout the report.

2 FR06/2021 The use of artificial intelligence and machine learning by market intermediaries and asset managers.

3 CR/01/2025 Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges.

international reports, and engagement with IOSCO Policy Committees, industry, and other stakeholders.

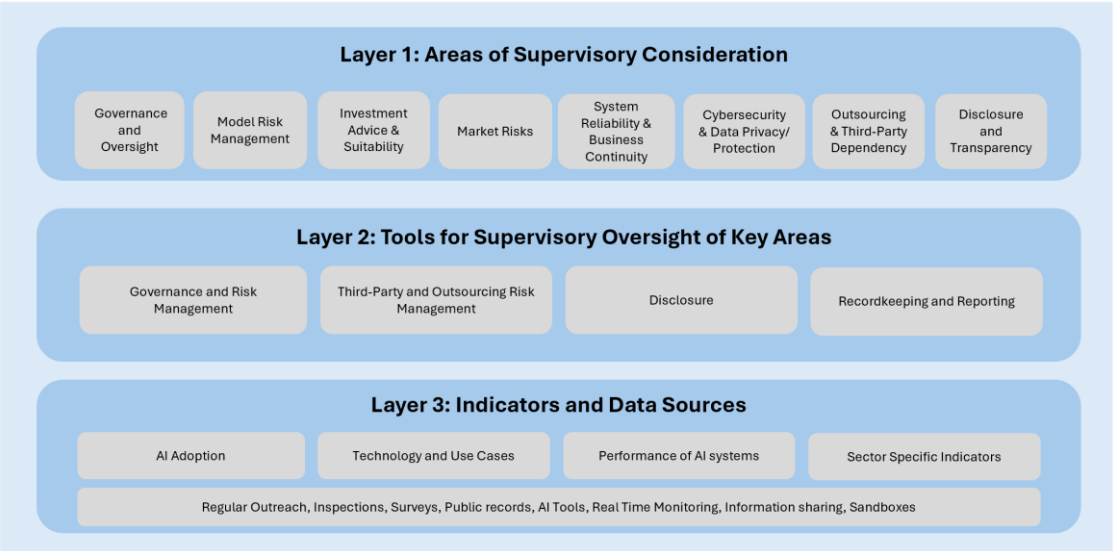
The report provides supervisors with a toolkit to support risk-based and proportionate supervision of AI use in capital markets through three complementary layers:

1. **Areas of supervisory consideration:** The first layer outlines areas of supervisory consideration relating to the use of AI systems in capital markets, building on the work conducted for previous IOSCO reports on AI. It provides supervisors with a framework for understanding which areas may warrant closer attention and where supervisory resources might be most effectively deployed.

2. **Tools for supervisory oversight of key areas:** This layer provides supervisors with more detailed tools to support evaluation across four areas of focus: (i) Governance and Risk Management; (ii) Third-party and Outsourcing Risk Management; (iii) Disclosure; and (iv) Recordkeeping and Reporting. It also includes practical examples of questions that supervisory authorities may find helpful for their dialogues with supervised firms and when planning examinations of supervised firms' AI use.

3. **Indicators and data sources:** The final layer equips supervisors with suggested indicators for monitoring AI adoption and use, alongside a range of engagement methods to gather relevant information. These include on-site inspections, targeted surveys, regular supervisory dialogue, requests for documentation and data, and other mechanisms that can support supervisory assessments of firms' AI systems.

Figure 1: Overview of the Supervisory toolkit



Supervisors may find it useful to refer to the accompanying document to this report, an extract of the standalone toolkit that can be readily utilized during supervisory activities, including on-site examinations and inspections. By consolidating the tools in a single, accessible format, this document is designed to serve as a hands-on resource that supervisors can refer to when assessing supervised firms' use of AI systems.[4]

Feedback and next steps

IOSCO welcomes feedback from stakeholders on the supervisory toolkit set out in this report. Respondents are invited to share their views on the tools in the report as well as any emerging practices in industry. Feedback may be submitted via a short survey by 26 June. IOSCO will take responses into account as it continues to develop its work on AI, including in the next phase of its review of emerging industry practices.

[4] This standalone document extracts the supervisory tools (tables) from Chapter 3 of the main report and compiles them in a concise format designed for practical use during examinations and inspections. It contains the same content as Chapter 3, but without the contextual analysis provided in the full report.

1: Introduction

1.1. Background

The rapid evolution of Artificial Intelligence (AI) has the potential to revolutionize capital markets through, among other things, enhancing efficiencies, improving processes, and detecting fraud. AI encompasses a number of technologies, including Machine Learning (ML), which has been embedded in financial products and services for some time, as well as more recent advancements like GenAI, which uses pre-trained models to synthesize or generate content. Emerging techniques involving Agentic AI, which are typically endowed with planning capabilities, long-term memory and access to tools, are expected to have profound impact on financial markets. As AI systems become increasingly embedded in capital markets, they bring opportunities but also corresponding risks inherent to their use. Given these evolving risks and the increasing adoption of AI systems across capital markets, there is a need for supervisory authorities to have practical tools to support effective oversight. This report responds to that need by providing IOSCO member authorities with a supervisory toolkit to support risk-based and proportionate oversight of AI systems used by supervised firms.

As the use of AI systems in financial products and services has expanded, IOSCO has undertaken a series of initiatives, including through its FTF and Artificial Intelligence Working Group (AIWG), [5] to assess the implications of these technologies for investor protection, market integrity, and financial stability. IOSCO's initiatives have evolved over time, reflecting both the development of AI systems and their increasing use by market participants in capital markets (see Annex II for more detail). IOSCO has also issued publications that, while not focused on AI, have discussed AI-related topics (see Annex III).

Since the publication of IOSCO's 2025 AI Report,[6] the use of AI systems in financial products and services has continued to operationalize and scale. Building on prior

[5] The AIWG is composed of the following IOSCO members: European Securities and Markets Authority (ESMA), Autorité des Marchés Financiers (AMF), Securities and Futures Commission (SFC), Securities and Exchange Board of India (SEBI), Central Bank of Ireland (CBI), Commissione Nazionale per le Società e la Borsa (CONSOB), Japan Financial Services Agency (JFSA), Ontario Securities Commission (OSC), Quebec Autorité des Marchés Financiers (QAMF), Monetary Authority of Singapore (MAS), Comisión Nacional del Mercado de Valores (CNMV), The Dutch Authority for the Financial Markets (AFM), Financial Conduct Authority (FCA), and the United States Securities and Exchange Commission (SEC).

[6] CR/01/2025 Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges.

advances using GenAI, firms have begun to shift from proofs of concept and pilot programs to broader integration of the AI use cases identified in the 2025 AI Report. These developments have been accompanied by heightened regulatory and supervisory attention, with emphasis on governance, disclosure, transparency, and explainability, as adoption of AI systems increases. International organizations and standard setting bodies, such as the FSB and others,[7] have continued to develop their respective work on AI use in the financial system.

This report builds on existing IOSCO work, alongside the work of other international organizations and supervisory authorities. It describes supervisory tools to assist IOSCO members in addressing the issues, risks, and challenges posed by the use of AI in capital markets.

Following the publication of this report, the AIWG will turn to a review of industry practices on the disclosure, recordkeeping and reporting, and governance of the use of AI systems in capital markets.[8]

[7] See, e.g., Monitoring Adoption of Artificial Intelligence and Related Vulnerabilities in the Financial Sector - Financial Stability Board, The IAIS publishes Application Paper on the supervision of artificial intelligence - International Association of Insurance Supervisors, BIS: Managing explanations: how regulators can address AI explainability, IMF: Artificial Intelligence, OECD: Full Report: Governing with Artificial Intelligence | OECD; and EIOPA Opinion on Artificial Intelligence Governance and Risk Management Opinion on Artificial Intelligence governance and risk management - European Insurance and Occupational Pensions Authority.

[8] In April 2026, Anthropic announced the launch of Claude Mythos, a frontier AI model reportedly representing a significant step-change from previous models. More broadly, recent developments in frontier AI models have highlighted autonomous capabilities, particularly the ability of such systems to independently discover, chain, and exploit security vulnerabilities (see, for example, Our evaluation of Claude Mythos Preview's cyber capabilities | AISI Work, Our evaluation of OpenAI's GPT-5.5 cyber capabilities). Developments in such capabilities may place greater emphasis on continuous risk assessment, integrated cyber and AI governance, and shorter remediation cycles. Anthropic has, in this context, opted for a gated release via Project Glasswing (Project Glasswing: Securing critical software for the AI era \ Anthropic). IOSCO is considering how these dynamics should be reflected as part of its forthcoming review of emerging industry practices on the disclosure, recordkeeping and reporting, and governance of the Use of AI Systems in capital markets.

1.2. Purpose

The aim of this supervisory toolkit for AI use in capital markets is to:

- Encourage risk-based and proportionate supervision from IOSCO Members in the context of AI use by supervised firms.
- Facilitate effective supervision by helping IOSCO Members identify key considerations as they approach AI oversight.
- Respond to technological advances by identifying ways in which supervisory approaches can evolve to keep pace with innovation, while at the same time preserving IOSCO's core objectives of market integrity, investor protection, and financial stability.
- Provide IOSCO Member supervisors with tools they may wish to incorporate as they conduct reviews of firms' use of AI.

This Report aims to be a resource for supervisors when approaching the oversight of AI systems used in capital markets but is not intended to be a binding or prescriptive guide.

1.3. Methodology

To inform this Report, the AIWG gathered evidence through:

(a) IOSCO member engagement: The AIWG conducted a survey of FTF members (the Survey) to understand how jurisdictions were approaching the evaluation and oversight of AI use in their supervised firms.[9] In addition, the AIWG engaged with IOSCO Policy Committees that have sectoral expertise[10] to understand developments in AI use in their respective sectors.

[9] 21 IOSCO members responded to the survey: ASIC (Australia), CVM (Brazil), OSC & QAMF (Canada), CSRC (China), AMF (France), BaFin (Germany), HCMC (Greece), SFC (Hong Kong), SEBI (India), CBI (Ireland), CONSOB (Italy), FSA (Japan), CMA (Saudi Arabia), MAS (Singapore), CNMV (Spain), FI (Sweden), FINMA (Switzerland), AFM (The Netherlands), FCA (United Kingdom), and SEC (United States).

[10] Committee on Regulation of Secondary Markets (Committee 2), Committee on Regulation of Market Intermediaries (Committee 3), and Committee on Investment Management (Committee 5).

(b) IOSCO's stakeholder engagement: The FTF hosted roundtables in Tokyo, Singapore, and New York, where participants presented and discussed with the FTF their perspectives on AI use in capital markets. Participants included supervised firms, other international organizations and standard setting bodies, academics, researchers, and AI technology service providers.

(c) Literature review: The AIWG reviewed publications by international organizations, academics, researchers, industry associations, market participants, and other stakeholders on AI use in capital markets, and its evaluation and oversight.

1.4. Definition of AI

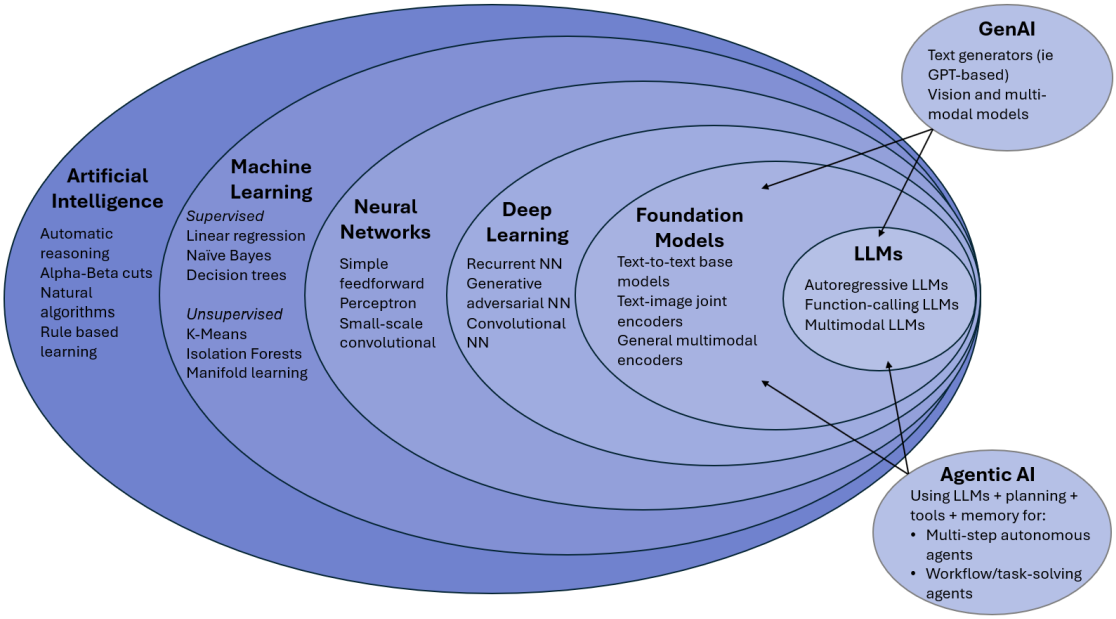
For the purpose of this report and consistent with the 2025 AI Report, IOSCO referred to the following OECD definition in its Survey:

> *"An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment".*[11]

Some members have defined the term "artificial intelligence" in developing their regulatory and supervisory approaches. These definitions generally refer to the use of machine-based systems simulating human intelligence, often including autonomy, adaptiveness, and inference capabilities, broadly aligning with the OECD definition. Given that definitions of specific types of AI technologies vary by jurisdiction and are likely to change as technology evolves, this report has not defined individual types of AI. Instead, it is more important to have a common understanding of the types of technologies discussed, informed by IOSCO's prior work in the 2025 AI Report. These types of technologies are illustrated in Figure 2.

[11] Updated 2024 definition as of: Explanatory memorandum on the updated OECD definition of an AI system | OECD.

Figure 2: Visual Representation of AI Technologies



Source: IOSCO AIWG, adapted from the 2025 AI Report[12]

1.5. Structure of report

Chapter 2 provides a brief update on developments in AI-related technologies and their applications in capital markets observed by IOSCO members. It introduces principles behind risk-based and proportionate supervision to frame the toolkit. It also provides an overview of primary risks, use cases, and types of AI (drawing on the March 2025 AI report).

Chapter 3 describes tools supervisory authorities may find useful to refer to in their oversight of AI use by supervised firms. These include tools for the monitoring of AI use in financial products and services, as well as key considerations and suggested questions across key potential risks.

Annex I details the results from the Survey, **Annex II** provides a timeline of IOSCO's work on AI, and **Annex III** outlines related IOSCO work.

[12] The categorization of different forms of AI is not always clear-cut. Boundaries can be blurry and definitions imprecise. Figure 2 is included as a representative visualisation to aid understanding of the relationships between different AI technologies, rather than as a definitive taxonomy. The diagram should be interpreted as illustrative of common categorizations and not as establishing rigid boundaries between AI system types.

2: Developments in AI use; risk-based and proportionate supervision

2.1 AI use case developments observed in capital markets

As noted in the 2025 AI Report, jurisdictions have reported that the use of AI in the financial sector is increasing, particularly for applications such as market analysis, forecasting and portfolio optimization, support, and risk management-related functions. The Survey results indicate relatively strong AI uptake overall in areas such as fraud detection and prevention, risk management, compliance, and client service and client engagement.

Some surveyed members reported that adoption of GenAI and Agentic AI techniques in financial products and services remains at a relatively early stage, characterized primarily by exploration and experimentation. While most activity appears concentrated in pilot projects and limited-scale deployments, Survey data indicates there is an increase in the speed and scale of AI development and implementation over the last two years, particularly using GenAI.[13]

The Survey data highlighted that regulators are increasingly aware of potential risks posed by Agentic AI, mainly if used in connection with retail financial products and services. While surveyed members indicated most uses of Agentic AI in financial products and services remain in the proof-of-concept phase, they cited

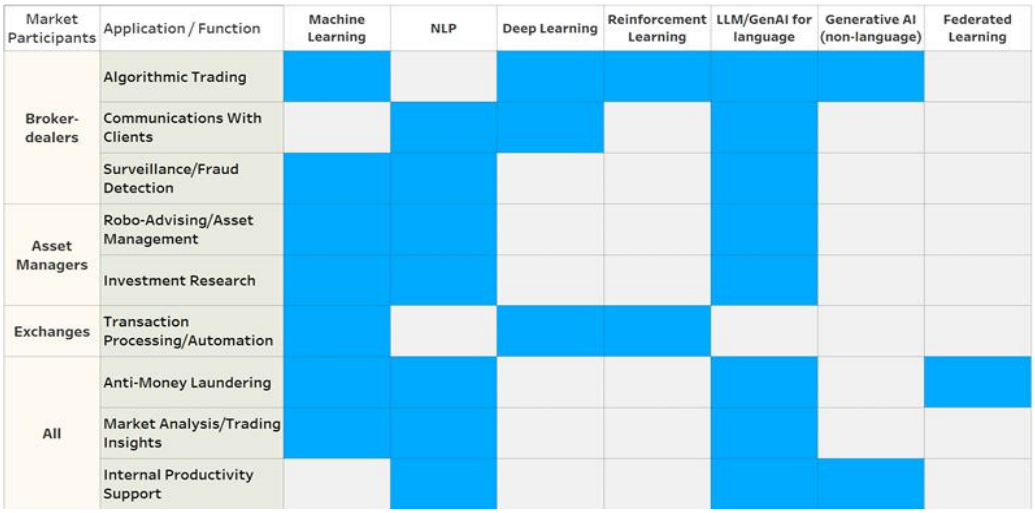
[13] While emerging, the increase in adoption of advanced AI is supported by other research: see, e.g., Financial services in the era of generative AI: Facilitating responsible adoption, Hong Kong Monetary Authority. Additionally, two responding IOSCO members noted experiments involving the use of AI technologies to automatically assess and oversee the performance of other AI systems ('AI as a judge'), as well as the use of AI to generate synthetic data for testing, model training, and privacy-preserving analytics and the integration of AI with blockchain technology for security, compliance automation, and transparency.

concerns with control over systems that incorporate Agentic AI techniques. Surveyed members indicated a need for increased focus on governance and accountability for such systems.

2.2 Drivers and potential benefits of AI adoption

In the 2025 AI Report, IOSCO identified types of AI technologies, applied to various AI uses, and grouped by market participant type, as shown in Figure 3 below.[14]

Figure 3: Type(s) of AI Technology Applied to Various AI Uses, and Grouped by Market Participant Type



Source: IOSCO AIWG, 2025 AI Report[15]

The emergence of more autonomous implementations of AI – sometimes referred to as Agentic AI – was predicted but not expanded on in detail in the 2025 AI Report. Given its potential future deployment in capital markets, this report includes

[14] See also IMF Technical Guidance Note: Regulatory Considerations Regarding Accelerated Use of AI in Securities Markets; December 2025; TNM 2025/16.

[15] The AI methods employed in these use cases encompass a range of techniques, including but not limited to: Machine Learning (an AI system designed to learn from experience without being explicitly programmed to do so); Deep Learning (an AI system involving neural networks (computing systems) with many layers of units, inspired by the structure of the human brain); Reinforcement Learning (an AI system that learns from receiving feedback, e.g., Q-learning); Natural Language Processing (techniques that enable computers to recognize, process, and generate text and speech, e.g., tokenization, TF-IDF, Latent Dirichlet Allocation, word2vec, GenAI for language tasks, and LLMs); Generative AI for non-language tasks (e.g., multi-modal systems); Federated Learning (technique of training models on decentralized data distributed across multiple devices).

high-level considerations to support supervisors in overseeing such advanced AI capabilities.

Since the publication of the 2025 AI Report, members have continued to observe increasing AI adoption driven by the following considerations:

(a) **Operational efficiency and cost reduction:** Surveyed members identified the desire for efficiency and cost reduction in the automation of back-office tasks and operational processes as a major motivator of AI adoption in capital markets, with firms seeking to reduce costs and streamline workflows. The Survey results included anecdotal reports that firms' use of AI for back-office functions has led to improvements in operational efficiency across financial institutions and that automation of routine processes reportedly has reduced manual workloads and operational costs; however, these reports do not quantify such improvements or reductions.

(b) **Data availability**: Surveyed members cited improved processes for collecting high-quality data to support the development and deployment of AI systems.

(c) **Client personalization and improved experience**: Surveyed members reported that AI is increasingly used, through chatbots and virtual assistants, to tailor services and enhance user experience. This includes AI use in systems that automate onboarding and support retail clients and those with smaller account balances.

(d) **Risk management and fraud detection**: Surveyed members reported that AI is being adopted for risk management, including credit risk assessment and fraud detection through analyzing large volumes of transactions to identify anomalies, suspicious patterns, and potential fraud and market manipulation. AI systems are also used to automate compliance checks and monitor adherence to regulations and support reporting. Some surveyed members noted that AI systems are being used to detect cyber threats and vulnerabilities.

2.3 Risks and areas of focus for supervisory consideration

The use of AI in capital markets presents opportunities and risks that supervisors must understand and address to protect investors, maintain market integrity, and preserve financial stability. Factors such as the rapid and unpredictable evolution of AI technologies and market dynamics add to the complexities of supervisory risk

assessment. Previous IOSCO reports detailed the potential risks of AI use in financial products and services, and these are briefly reiterated below.

Table 1: Key findings of past reports

Past Report	Uses Cases	Potential Risks
2021 AI Report	Found that, at the time, market intermediaries were deploying AI for: advisory and support services; risk management; client identification and monitoring; selection of trading algorithm; and asset management and portfolio management.	Identified potential risks that may arise in relation to the development, testing, and deployment of AI, in the areas of governance and oversight; algorithm development, testing, and ongoing monitoring; data quality and bias; transparency and explainability; outsourcing; and ethical concerns.
2025 AI Report	Found that, at the time, firms were increasingly using AI systems to support decision-making processes in applications and functions such as robo-advising, algorithmic trading, investment research, and sentiment analysis. Firms were also using AI to enhance surveillance and compliance functions, particularly in AML and CFT measures (transaction monitoring). Firms were using or considering LLMs and GenAI uses in: supporting internal operations and processes through task automation; enhancing communication via chatbots; and improving risk management functions.	Identified the most cited potential risk areas as malicious uses; model and data considerations; concentration, outsourcing, and third-party dependency; and interactions between humans and AI. Also flagged areas for future monitoring and warned that data and knowledge gaps can inhibit full risk assessment, particularly gaps around how to measure and manage "macro" risks, i.e., the impact of aggregate firm-level conduct to system-wide stability.

Together, these reports reflect both the persistence of certain risks associated with AI broadly, and the emergence of new challenges associated with recent advancements in AI technologies, such as GenAI and Agentic AI systems.

As detailed in the 2025 AI Report, any evaluation of the risks of AI use in financial products and services is highly contextual and will depend upon a number of factors, including the use case for the AI system, the choice of its design and what particular technologies are employed, how it is deployed, and how the environment in which it is used changes over time. The degree to which there is human oversight

for the AI system and its operation, and its potential impact on investors and the markets, also factor into its risk profile. Additionally, AI use raises a continuum of supervisory issues along the stages of the AI system's lifecycle.

Also, as explained in the 2025 AI Report, developments in AI technologies can result in new types of risks or existing risks manifesting in new ways. For example, the use of GenAI in financial products and services can amplify and add to risks of more traditional AI systems, requiring enhanced supervisory approaches. Due to their non-deterministic nature and technological complexity, GenAI systems may be difficult to predict, evaluate, understand, explain, and test. One such risk in the GenAI context arises from hallucinations, as explained in Box 1 below.

Box 1: Hallucination risk arising from use of AI systems

One pertinent concern raised by IOSCO members is that of so-called "hallucinations" (or "confabulations"), which refers to the generation by an AI system of outputs that appear plausible but are factually or logically incorrect or fabricated. Hallucination is a risk in GenAI systems, which use LLMs relying on probabilistic word generation and which can produce inconsistent outputs even when given identical or similar inputs. While hallucinations may not be intended, they can represent a critical risk for financial services, where trust and credibility are paramount, as hallucinated outputs incorporated into financial products and services can in certain use cases lead to investor and market harm.[16]

Approaches being explored to address the risk of hallucinations include:

(a) Grounding and guard-railing techniques: Technical grounding and guard-railing measures can reinforce the factual accuracy of responses. These techniques include Retrieval-Augmented Generation (RAG), Chain- of-Verification (CoVe), and Multi-Agent Debate,[17] among other

[16] Supervisors may also find it useful to be aware of what remedial measures can be taken in their respective jurisdictions to address investor loss from the use and reliance on inaccurate outputs from AI systems. Some IOSCO members are considering reciprocal compensation frameworks, however, jurisdiction-level discussions on such frameworks remain ongoing and are not considered further in this report, which focuses on supervisory tools.

[17] RAG: a mechanism that searches and retrieves curated knowledge bases or documents before generation a response. CoVe: a technique designed to reduce hallucinations in large language models by introducing an additional verification step after the initial reasoning or generation process and forcing the model to verify its own draft before it responds. Multi-agent debate: Multiple LLMs or agents engage in mutual critique, discussion, and voting.

techniques. It should be noted that these techniques have limitations and are unlikely to fully eliminate the risk of hallucinations.

(b) Human oversight: Ensuring that a human overseer is part of the AI system at the appropriate point(s) in the AI system's lifecycle is widely considered as an effective measure. However, it is pointed out that human overseers may be susceptible to automation bias, may tend to over-rely on AI, and may be unable to adequately oversee certain processes in an AI system. Robust governance and risk management at the firm level should consider how human oversight can be effectively implemented, including whether the human overseer has the requisite capabilities and knowledge of the system's design and limitations to exercise effective oversight and whether there is sufficient accountability for human oversight.

(c) Education: Generally, firms are responsible for the financial products and services they provide, regardless of whether they use AI or some other technology. Thus, the primary responsibility for addressing hallucination risk of GenAI used in financial products and services is that of the firm deploying the GenAI system. Alongside reasonable steps by firms, including appropriate disclosures, investor education on AI risks can improve investors' and users' awareness of the risks involving AI systems, including that of hallucinations. [18]

As anticipated in the 2025 AI Report, Agentic AI systems and techniques are emerging. These systems typically consist of various components, including pre-trained models, prompt registries, tools, data sources, memory, and a multitude of agents for different tasks (e.g., for data retrieval, incident classification, and quality assurance tasks).[19] In deployment, Agentic AI systems are typically equipped with planning capabilities, have long-term memory, and can access external tools, systems, or other agents. Consequently, Agentic AI systems typically can execute computer code, use the internet, or perform real-world tasks.

The risks from the use of GenAI and LLMs, such as hallucinations and other data and model considerations (outlined in the 2025 AI Report), are equally relevant for Agentic AI. But, while Agentic AI has the potential for myriad applications, the complexities of Agentic AI further compound the complexities of risk assessment

[18] In March 2026 IOSCO launched a TechSprint project to create a globally replicable tool to educate retail investors on AI related risks, including hallucination risk: IOSCO announces Call for Applications for its first TechSprint on Investor Education in the Age of Artificial Intelligence (AI).

[19] See Deloitte (Dec. 2024): The cognitive leap. How to reimagine work with AI agents.

and, if not addressed, can introduce risks that can have significant consequences in financial markets. Agentic AI systems can have access to sensitive data, can become exposed to malicious external content, and can have the ability to communicate externally. If an Agentic AI system is not configured properly, the combination of these features could result in data compromise, exfiltration, and operational and cybersecurity issues. Vulnerabilities caused by poorly designed prompts, inadequate safeguards or insufficient testing and monitoring of agent interactions could lead to poor performance and failures in Agentic AI systems.

Since Agentic AI uses more components than less complex GenAI systems, the potential interplay between the components could create the risk of unexpected emergent behaviors or the potential for cascading impacts or failures across interconnected systems. More specifically, increased complexity and opacity, coupled with increasingly automated workflows, can create the potential for unpredictable or unwanted behaviors. For example, an AI agent may take undesired steps to pursue its goal, such as engaging in collusive behaviors with other components or systems, or its goal may become misaligned from the one for which it had been deployed. Furthermore, there are challenges with detecting and addressing unwanted activity in Agentic AI systems due to their complexity and opacity.

Agentic AI may make supervisory oversight more challenging as this technology evolves and is incorporated into financial products and services, suggesting the need for a greater understanding of how governance and risk management systems can effectively be implemented to oversee the use of Agentic AI in financial products and services.

At the same time, technological developments may also advance efforts in explaining and interpreting AI systems or improving the accuracy of results. Market participants may also develop methods to validate AI system outputs in cases where understanding or interpreting an AI system is not feasible. In light of these developments, supervisors may wish to continue monitoring technological developments in administering their supervisory frameworks and processes.

2.4 Risk-based supervision and proportionality

Risk-based supervision represents a foundational framework for effective oversight in the financial sector and is equally applicable to oversight of the use of AI in financial products and services. This approach recognizes that different use-cases of AI systems pose different levels of risk and that supervisory expectations should be higher for use cases carrying higher levels of risk, particularly in respect of investor protection, market integrity, and financial stability. The principle of proportionality is central to a risk-based supervisory framework, ensuring that supervisory resources are deployed in a way that accounts for the level of risk

presented by a given activity, based on the potential impact on investor protection, market integrity, and financial stability.

These concepts can guide supervisors in maintaining appropriate oversight proportionate to the intensity and scope of the risks that AI systems can generate, while permitting innovation to occur, thereby balancing investor and market protection with fostering responsible advancements in technology.

Box 2: Risk-Based and Proportionate Supervision: Considerations for AI Systems

In applying a risk-based and proportionate approach to supervision of AI use in financial products and services, supervisors could begin with an assessment of risks considering a number of factors.

For example, supervisors could evaluate risks by examining such factors as the nature and complexity of the relevant AI system, the level of human oversight of the AI system, and its potential implications for investors and the broader market.

(a) **Nature and complexity of AI system:** Supervisors could consider the fact that different types of AI systems pose different types of risk. For example:

(i) More complex or difficult to interpret AI systems may limit the ability to identify unintended behaviour and to diagnose the root causes of incidents.

(ii) AI systems that process near real-time updates in data may not perform well if the data differs from that on which the system was trained.

(iii) GenAI systems may generate fundamentally incorrect outputs ("confabulations" or "hallucinations")[20] that are convincing but inaccurate or unsuitable.

(b) **Human Oversight:** Supervisors could consider the extent of human oversight of the AI system and, relatedly, dependencies on AI components and third-party service providers. The extent of human oversight on the AI system is relevant to assessing the risk that an issue with the AI system results in actual harm, before the issue can be

[20] See Box 1 for an expanded explanation on "hallucination risk" arising from AI systems.

detected and mitigated. The various levels of human oversight of an AI system can be classified as follows:[21]

(i) *Human-in-control:* The AI system cannot act on its recommendations or output. The human uses or disregards the AI system's recommendations or output at the human's discretion.

(ii) *Human-in-the-loop:* The AI system evaluates input and acts upon its recommendations or output if the human approves.

(iii) *Human-on-the-loop:* The AI system evaluates input and acts upon its recommendations or output unless the human disapproves.

(iv) *Human-out-of-the-loop:* The system evaluates input and acts upon its recommendations or output without human involvement.

(c) **Potential Impact of AI Use on Clients or the Firm:** Supervisors could consider the potential impact of the AI system as relevant to assessing the severity of harm that could be caused by an incident.

(i) Factors relevant to assessing the impact on clients could include threats to investor protection, such as threats to suitability, privacy, data protection, and non-discrimination (i.e., where financial inclusion is at stake); the volume and type of investors affected; and the ability and willingness to promptly identify and address investor harm (including whether there are effective redress mechanisms).

(ii) Factors relevant to assessing the impact on a firm-could include whether, either external or internal-facing, AI systems could disrupt core business activities, cause a system to fail, or result in adverse operational, legal and/or reputational consequences.

(iii) The source of the potential impact could also be assessed whether from a vulnerability in the system (model or data), the inappropriate use of a system (misuse or cyber incident), or ineffectual risk mitigation (lapse in human oversight).

In analysing AI use within a firm, supervisors may also consider whether the firm has a governance and risk management framework calibrated to the nature of use cases at the firm and their potential outcomes, including the necessary knowledge and resources for adequate testing, maintaining, and

[21] OECD Framework for the Classification of AI Systems, p. 53. Such systems have also been classified as Assisted; Augmented; Adaptive; and Autonomous, respectively.

monitoring AI systems; measures around data quality; and the appropriate level of transparency for AI use for stakeholders.

The factors outlined in Box 2 are illustrative, and not exhaustive. A comprehensive impact assessment or risk mapping helps enable supervisors to classify AI systems along a spectrum of risk and proportionality, and to then assess what supervisory measures should be taken to address the AI system. Analysis of appropriate factors for a particular AI system should result in a risk-based and proportionate approach, with higher-risk/higher-impact systems subject to more stringent examination and inspection, and lower-risk/lower-impact systems subject to more streamlined processes.

In this respect, larger, systemically important institutions may face heightened supervisory expectations even for medium or low-risk AI applications, given their potential market-wide impact. Relatedly, as regards concentration risk and outsourcing, in cases where multiple institutions rely on the same third-party AI providers, supervisors could consider assessing systemic vulnerabilities, recognizing that AI failures in systemically important firms or shared infrastructure could have cascading effects.

3: Supervisory toolkit

This chapter is the primary focus of the Report and aims to provide supervisory authorities with practical examples of non-binding, non-prescriptive tools to support them in their oversight of AI system use in capital markets' financial products and services. The tools are not exhaustive but provide supervisors with a structured starting point and a flexible framework to identify risks and tailor their oversight approaches to the specific characteristics and contexts of AI deployment within their capital markets.

Table 2 provides an overview of the primary areas of supervisory consideration, to promote responsible deployment of AI systems that advances benefits while mitigating associated risks. The table highlights where there is overlap with the measures recommended in the 2021 AI Report.[22] For reference, Box 3 outlines the measures from that report.

Table 2 is followed by more specific tools, aimed at supporting supervisors in their oversight of key areas including: Governance and risk management; Third party risk management; Disclosure; and Recordkeeping, reporting, and monitoring of AI use.

Box 3: Measures outlined in March 2021 AI Report

Measure 1: Regulators should consider requiring firms to have designated senior management responsible for the oversight of the development, testing, deployment, monitoring and controls of AI and ML. This includes a documented internal governance framework, with clear lines of accountability. Senior Management should designate an appropriately senior individual (or groups of individuals), with the relevant skill set and knowledge to sign off on initial deployment and substantial updates of the technology.

Measure 2: Regulators should require firms to adequately test and monitor the algorithms to validate the results of an AI and ML technique on a continuous basis. The testing should be conducted in an environment that is segregated from the live environment prior to deployment to ensure that AI and ML: (a) behave as expected in stressed and unstressed market conditions; and (b) operate in a way that complies with regulatory obligations.

[22] FR06/2021 The use of artificial intelligence and machine learning by market intermediaries and asset managers.

Measure 3: Regulators should require firms to have the adequate skills, expertise and experience to develop, test, deploy, monitor and oversee the controls over the AI and ML that the firm utilises. Compliance and risk management functions should be able to understand and challenge the algorithms that are produced and conduct due diligence on any third-party provider, including on the level of knowledge, expertise and experience present.

Measure 4: Regulators should require firms to understand their reliance and manage their relationship with third-party providers, including monitoring their performance and conducting oversight. To ensure adequate accountability, firms should have a clear service level agreement and contract in place clarifying the scope of the outsourced functions and the responsibility of the service provider. This agreement should contain clear performance indicators and should also clearly determine rights and remedies for poor performance.

Measure 5: Regulators should consider what level of disclosure of the use of AI and ML is required by firms, including: (a) Regulators should consider requiring firms to disclose meaningful information to customers and clients around their use of AI and ML that impact client outcomes. (b) Regulators should consider what type of information they may require from firms using AI and ML to ensure they can have appropriate oversight of those firms.

Measure 6: Regulators should consider requiring firms to have appropriate controls in place to ensure that the data that the performance of the AI and ML is dependent on is of sufficient quality to prevent biases and sufficiently broad for a well-founded application of AI and ML.

Table 2: Areas of supervisory consideration

Area	IOSCO 2021 Report Measures (see Box 3 above)	Potential Supervisory Concerns	Supporting Evidence for Review
AI Governance & Oversight (see Table 3 for more detailed considerations)	Measures 1 and 3	• Lack of Board and senior management oversight or accountability; inadequate reporting to Board and senior management. • Lack of documented internal AI governance and risk management framework and broader AI policies and procedures, including IT and data governance framework. • No AI inventory or classification to identify AI use cases. • Insufficient approval processes. • Inadequate training and ongoing education for Board, senior management and staff exercising control functions. • Lack of understanding and knowledge of AI system design. • Lack of clear and documented roles and responsibilities for developers, deployers, and users.	• AI governance policies & procedures. • AI risk management framework. • AI inventory/registry, including technical documentation/description of AI systems. • Governance committee documents. • Organization charts. • Training programs, materials, and records relating to AI systems and usage. • Human oversight, policies & procedures in place for appropriate and proportionate human intervention/interruption in the operation of the AI system. • Documentation of staff qualifications, including relevant certifications, competency assessments, and continuing education plans.
Model Risk Management (see Table 3 for more detailed considerations)	Measures 2, 3 and 6	• Inadequate model testing, including back testing, stress testing, testing for bias and model drift, and underlying data testing. • Inadequate ongoing performance monitoring.	• Model validation and testing policies & procedures and reports, including pre-deployment, post-deployment and ongoing validation, monitoring and testing, including documentation and logs.

Area	IOSCO 2021 Report Measures (see Box 3 above)	Potential Supervisory Concerns	Supporting Evidence for Review
		• No independent validation of model performance. • No alert mechanisms for anomaly detection. • No methodology for AI system suspension where anomalies are detected.	• Model performance monitoring policies & procedures and reporting, including performance and bias. • Model change management policies & procedures and logs. • Independent AI model validation and testing reports. • Model performance anomaly detection alerts and processes.
Investment Advice & Suitability (see Table 5 for more detailed considerations)	Measure 6	• Lack of suitability of AI-generated recommendations. • Failure to consider investor circumstances. • Inadequate human oversight. • Bias in AI systems, including logic and prompts. • Conversational interface advertising (i.e., embedded sponsored content within an AI-driven advisory conversation). • Claims that are misleading or overstate AI capabilities ("AI-Washing").	• Client profile and input data. • Suitability policies and procedures addressing client profile, including risk tolerance and investment objectives. • Policies and procedures addressing the availability of product or service offerings for investment recommendations by any AI system. • AI recommendation logs, including outputs and supporting data. • Human review and override documentation. • Conflict identification and management policies & procedures. • Investor complaints. • Client disclosures.

Area	IOSCO 2021 Report Measures (see Box 3 above)	Potential Supervisory Concerns	Supporting Evidence for Review
Market Risks (see Table 3 for more detailed considerations)	Measure 2	• AI use amplifying market volatility. • Flash crashes from AI-driven trading. • Herding behavior/correlation /collusion. • Liquidity issues. • Other sources of systemic risk.	• Market risk policies & procedures. • Stress testing policies & procedures. • Circuit breaker or kill switch policies & procedures. • Volatility and liquidity management plans. • Emergency policies & procedures.
System Reliability & Business Continuity Planning	Measures 2 and 6	• AI system failures. • No backup or recovery procedures. • Single points of failure. • Inadequate business continuity. • Service disruptions.	• Operational resilience framework, encompassing both business and cybersecurity aspects. • Business continuity and disaster recovery planning and testing, including plans for AI service disruptions and system outages. • Backup and recovery policies & procedures, including safeguards to protect client records and other sensitive information. • System availability and performance monitoring, including related reporting and service level agreement compliance.

Area	IOSCO 2021 Report Measures (see Box 3 above)	Potential Supervisory Concerns	Supporting Evidence for Review
Cybersecurity & Data Privacy/Protection	Not covered explicitly in 2021 AI Report.	• AI system attacks/breaches. • Client data exposure or exfiltration. • Model theft or manipulation (poisoning). • Inadequate access controls. • Other data leakage. • Use of AI for advanced cyberattack, i.e., social engineering, deepfakes, identity theft.	• Cybersecurity policies & procedures and standards, including where relevant the cloud infrastructure associated with AI systems used. • Penetration test framework. • Identity and access management controls for AI systems and data. • Incident response policies & procedures. • Privacy impact and data privacy assessments. • Security audit reports, logs, and penetration test results. • Continuous training of users.
Outsourcing & Third-Party Dependencies (see Table 4 for more detailed considerations)	Measures 3 and 4	• Third-party AI vendor risks, including data access and privacy risks, and cybersecurity risks. • Inadequate due diligence. • Poor contract terms. • Inadequate monitoring of the third-party Provider. • Vendor concentration and dependency risk. • Service provider failures. • Lack of technical skills involved in procurement process.	• Vendor selection, due diligence (including the level of knowledge, expertise and experience on AI systems), and contract terms, including notice and exit provisions, service level requirements, and data protection obligations. • Ongoing monitoring and oversight procedures, including performance metrics, issue escalation policies, and remedies for poor performance. • Third-party validation and assessments of AI systems, including the identification of cross-market dependencies and potential single points of failures affecting multiple entities.

Area	IOSCO 2021 Report Measures (see Box 3 above)	Potential Supervisory Concerns	Supporting Evidence for Review
		• Inappropriate cross-jurisdictional data transfer.	• Where contract terms are general/not tailored to the firm, an assessment of the risks associated with using the firm's services and the firm's own processes to manage those risks.
Disclosure & Transparency (see Table 5 for more detailed considerations)	Measure 5	• Inadequate AI disclosure to investors. • Claims that are misleading or overstate AI capabilities (AI-Washing). • Hidden AI system usage for investor facing services. • Lack of transparency about limitations. • Insufficient disclosure of use of third-party AI services. • Over-reliance on third-party providers' disclosures.	• Client agreements, including account information, acknowledgements, and marketing materials relating to AI usage. • Disclosures outlining material risks and impacts on investors. • Disclosure of incidents where adoption of AI systems has raised regulatory, ethical or legal issues. • Client communications relating to AI usage, including policies & procedures for monitoring such communications for accuracy. • Policies & procedures for reviewing and updating client disclosures to ensure AI-related disclosures are accurate and up-to-date.
Recordkeeping & Audit Trail (see Table 6 for more detailed considerations)	Not covered explicitly in 2021 AI Report.	• Inadequate AI system records. • Lack of explainability of AI logic. • Missing audit trails or logs. • Insufficient communication with supervisors through required regulatory reporting.	• Recordkeeping policies & procedures for AI systems and data, including data with a third-party or external provider. • AI inventories. • Recordkeeping of AI-generated outcomes and how AI systems generate outputs.

Area	IOSCO 2021 Report Measures (see Box 3 above)	Potential Supervisory Concerns	Supporting Evidence for Review
		• Insufficient documentation. • Lack of AI system oversight throughout the lifecycle.	• Compliance policies & procedures for adherence to laws, regulations, and internal standards. • AI system decision and usage logs and audit trail documentation, including inputs, outputs, and AI system logic. • Regulatory filings and supporting documentation. • Incident reporting.

3.1 Governance and risk management

Robust governance and risk management frameworks are essential to ensuring responsible and resilient deployment of AI systems into financial products and services. These frameworks provide the structural foundation for identifying, assessing, and mitigating risks associated with AI technologies, including regulatory, operational, and other risks. Strong governance and risk management practices not only safeguard investor protection and market integrity but also enhance investor confidence by demonstrating that AI systems are deployed in a manner consistent with principles of fairness, explainability, transparency, interpretability, and operational resilience.[23]

Effective risk management frameworks typically provide for investor and market protections and encompass the entire AI lifecycle – from design and development of AI systems to deployment and ongoing operation, and to retirement – supported by documented policies and procedures. This includes controls for use case design as well as model, data, and third-party risks. The risk management of AI systems typically includes measures to promote transparency, reliability, robustness, resilience, fairness, security, safety, and privacy. There is typically a clear assignment of roles and responsibilities and measures for appropriate human oversight (e.g., human-in-the-loop), especially where outcomes can impact investors and the markets. Firms may also have escalation protocols for AI-related incidents and maintain comprehensive audit trails to enable supervisory review as well as adequate reporting to senior management and boards.

Effective governance also promotes accountability by establishing clear roles and responsibilities for senior management and boards, ensuring that decision-making and oversight mechanisms are proportionate to the materiality and complexity of AI use cases. As AI adoption accelerates, these frameworks will likely need to evolve to address emerging risks or heightened risks, including for example, concentrations in and dependencies on third-party providers, adversarial attacks,

[23] Transparency, explainability, and interpretability are distinct characteristics that support each other. Transparency reflects the extent to which information about an AI system and its outputs is available to individuals interacting with such a system. Explainability refers to a representation of the mechanisms underlying AI systems' operation, whereas interpretability refers to the meaning of AI systems' output in the context of their designed functional purposes. Transparency can answer the question of 'what happened' in the system. Explainability can answer the question of 'how' a decision was made in the system. Interpretability can answer the question of 'why' a decision was made by the system and its meaning or context to the user. AI Risk Management Framework, pp. 15-17; see also OECD AI Principle 1.3. on transparency and explainability, https://oecd.ai/en/dashboards/ai-principles/P7.

and unintended market impacts, ensuring that financial institutions remain agile and accountable in a rapidly changing technological landscape.

Common elements of governance and risk management frameworks

The 2025 AI Report described steps market participants had been taking at the time of that report to manage risks, and govern internal development, deployment, and maintenance of AI systems.[24] Responses to the Survey update that information, and commonly identify the following elements as part of observed governance and risk management frameworks related to AI use:

(a) **Centralized Governance Structure for AI Oversight to Avoid Gaps in AI Governance and Risk Management.** These centralized structures at the firm level can serve as key platforms for coordinating governance and oversight of AI usage across the organization and can play an important role in addressing emerging challenges and potential gaps in risk management as the AI landscape evolves. These structures could take the form of cross-functional, executive-level committees with clear responsibility and decision-making authority over AI use and governance. They also can extend to board-level engagement to adequately fulfill the board's risk oversight responsibilities.

(b) **Firm-wide AI Policy Framework:** Policies and procedures that are relevant to AI can be applied across a firm to promote consistent principles and standards for AI and clear roles and responsibilities across the firm. Baseline principles and standards can be supplemented or enhanced for higher risk AI use cases.

(c) **AI System Risk Management, including validation and testing, across the lifecycle for AI systems:** Appropriate risk management, including validation, testing, and monitoring, generally applies for all forms of AI systems. Notably, a key distinction for GenAI systems is that system behavior cannot be fully specified or anticipated in advance of deployment; whereas traditional software behaves according to rules that developers specify and understand, enabling strong guarantees about how the system will perform across inputs and over time. GenAI systems, by contrast, derive their behavior from, e.g., data, training, optimization dynamics, and emergent internal representations. GenAI relies on underlying models, which are trained on data, and learning a task from data rather than using explicit programming introduces a fundamentally

[24] See Section V of the March 2025 AI Report: CR/01/2025 Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges.

different uncertainty profile for such systems. As a result, a GenAI system loses many of the guarantees that make rules-based systems predictable. To mitigate challenges related to the risks of such systems, firms can maintain comprehensive documentation; conduct formal and independent reviews; validate systems prior to and after deployment (e.g., using stress testing and scenario analysis); conduct ongoing performance monitoring to detect issues such as drift, degradation or bias of underlying models; and benchmark outputs against industry standards to assess quality and consistency.

(d) **AI Inventory Management:** Firms can maintain an AI inventory with a comprehensive record of the various AI systems used in the firm (both third-party and internally developed), and how these AI systems are used in the firm. The inventory would capture such key attributes as, e.g., the AI system's purpose and description, scope of use, data and model usage, and upstream and downstream dependencies. The firm can develop a methodology to assess the potential risks and impacts of an AI system on an ongoing basis.

(e) **Capability Building and Training:** Firms can establish plans to upskill their staff, senior management and Board, and develop AI training that facilitates staff in leveraging and using AI in an effective and responsible manner. As seen in practice in big financial firms, different types and levels of staff might be subject to different levels of training, depending on the roles and functions they perform. Some firms have set up AI Centers of Excellence to drive innovation, promote good practices, and build AI capabilities.

(f) **Generative AI Specific Practices:** For GenAI, emerging practices include focusing on the development of key enabling modules or services that can be reused across multiple use cases and establishing pilot and experimentation frameworks with clear policies and procedures for GenAI pilots that are typically bound by time and user limits and often occur in a contained environment or with non-sensitive data. Similarly enhanced processes may also be appropriate for use of other AI systems using emerging technology and techniques like Agentic AI.

Supervisory toolkit

Table 3 outlines various areas, including those highlighted in the Survey, for consideration by supervisory authorities in the oversight of governance and risk management frameworks, including key considerations and examples of questions that supervisors may find useful.

Table 3: Evaluating governance and risk management of AI use

Area	Considerations	Examples of Questions	Information Sources
1. Oversight from Board or Governing Body	• Does the Board set strategic objectives and risk appetite for AI use, and does it set AI policies and procedures? • Does the Board receive regular management information on areas such as the design, implementation, and use of AI systems, governance and risk management, and measures to address investor protection and market integrity?	• Has the Board defined an AI strategy and risk appetite and communicated that within the organization? • How does the Board ensure it possesses sufficient understanding of AI risks and opportunities and how does it promote a corporate culture that prioritizes ethical, fair, and responsible AI use across the organization? • How does the Board ensure that AI risks, where material, are explicitly addressed within the firm's overall risk appetite and management framework, including the setting of appropriate qualitative statements and quantitative measures or limits? • How does the Board ensure that the firm's approach, risk management framework, roles and responsibilities, capabilities and culture for risk management of AI use are regularly reviewed to keep pace with newer AI developments, as well as changes in the firm's risk profile and business strategies? • Does the Board approve the risk management framework, the operational resilience framework, and the outsourcing due diligence framework, and related material modifications? • How does the Board ensure that clear accountability mechanisms are in place when it comes to AI and its risks? • What processes are in place for the Board to independently review and challenge management's assessment of AI system risks, including validation results, AI incident investigations, and the adequacy of mitigation actions? • What reporting is provided to the Board on the use of AI models or systems? How frequently is this provided?	• The firm's AI strategy, risk appetite, AI policies and procedures or broader governance and risk management framework. • Board minutes, management information packs, board reports, key performance indicators. • Evidence of board training on AI governance and risk management.
2. Senior Management Responsibilities	• Is accountability for AI assigned appropriately to	• Describe the firm's framework for human oversight of AI systems, including clear allocation of roles and responsibilities, decision rights, and escalation paths for AI-driven outcomes.	• The firm's AI policies and procedures or

Area	Considerations	Examples of Questions	Information Sources
	senior management? • Is senior management responsible for ensuring the effective implementation of AI-related governance and risk management policies and procedures, and regularly reviewing their effectiveness?	• Describe the processes, and senior management's roles and responsibilities within these processes, for the following: ○ Introduction and implementation of AI systems. ○ Coordination and accountability for AI-related risk management across the firm. ○ Review of the use of the AI system for compliance with legal and regulatory requirements. ○ Internal escalation process for managing material AI risks and exceptions, such as incidents or breaches of risk thresholds, and ensuring appropriate and timely actions are taken. ○ Updating the Board on material AI risk issues in a timely manner. ○ Ensuring the necessary competence of personnel and allocating adequate resources (such as human, technological, financial resources) for effective AI risk management, including appropriate training and capacity building. ○ Establishing controls over use of AI by staff, including third-party tools or AI systems used without prior authorization. ○ Establishing adequate controls around third-party vendors.	broader governance and risk management framework • Evidence of risks that have been escalated to senior management and/or the Board • Organizational charts and role/responsibility assignment matrices that delineate roles/responsibilities
3. AI risk management systems, policies and procedures	• Does the firm implement a formal AI risk management framework, and how does this risk management framework incorporate considerations of the key sources of risk in AI systems, e.g., hallucination	• What is the firm's stated risk appetite for AI systems (including GenAI and Agentic systems)? Where is risk appetite formalized within the firm internal documentation? • How does the firm maintain a complete AI inventory, including vendor information, usage, risks and goals of each AI system or model, and govern their use according to the assessed risk appetite? • Does the firm have an assessment methodology to evaluate the risk materiality of an AI use case, system or model based on the nature of its business? Is this risk assessment done on a regular basis? Has an impact assessment been undertaken?	• The firm's AI policies and procedures or broader governance and risk management framework • Lifecycle standard operating procedures, system

Area	Considerations	Examples of Questions	Information Sources
	risk, explainability and transparency, conflict of interests and bias? • Does the firm determine appropriate human-oversight configurations, define intervention thresholds, support contestability and redress where appropriate, and guard against automation bias? • Does the firm operate end-to-end AI system lifecycle controls (from ideation to retirement)?	• Has the firm developed indicators to monitor performance, and has the Board/senior management approved their use? • For a material AI use case, what human oversight measures are in place? Who can intervene or override and how quickly? How often do the interventions/overrides take place? How does the firm back-test if such overrides are appropriate? • What policies and procedures are in place to identify, address or mitigate material conflicts of interest in the use of AI systems? • How does the firm monitor for and guard against automation bias, including through training, awareness information programs and/or controls that encourage critical assessment of AI outputs by staff? • How is AI risk management integrated into the firm's overall risk management framework? • Explain in detail the governance of the AI system lifecycle for a high-impact AI use case, from ideation to retirement. • Does an Internal Audit function conduct reviews concerning the use of AI within the organization?	development standards, validation/test plans and go-live approvals. • Internal audit reporting
4. Data governance	• Does the firm, through its policies and procedures ensure data used for AI systems is accurate, complete, representative, timely and relevant	• How does the firm ensure the quality, representativeness, and appropriateness of data used in AI systems, including steps taken to identify and mitigate potential biases? • What governance structures exist around data ownership and accountability? • How does the firm ensure transparency and traceability of data used in AI systems? • What processes does the firm have in place to ensure that the use of personal or sensitive data in AI systems complies with applicable regulations?	• The firm's data governance framework, data management policy, or enterprise-wide data strategy

Area	Considerations	Examples of Questions	Information Sources
	for the intended purpose? • Does the firm conduct appropriate data collection, preparation, and maintenance in the use of AI systems?	• How is data and system access provided AI systems to protect material non-public information and respect information barriers? • How does the firm validate the suitability of external or third-party data sources, and what due-diligence processes are in place? • How does the firm monitor and manage data degradation? • What mechanisms are in place to detect and remediate data incidents, including inaccuracies, loss, corruption, or unauthorized access?	• Records of data quality assessments, profiling exercises, and remediation logs
5. Risk management of Advanced AI systems	• Does the firm clarify and understand distinctions among different forms of AI within its internal risk management and operational resilience documentation, particularly for higher risk advanced AI systems, and manage and oversee those systems commensurate with the risk involved?	• Does the firm distinguish between different AI systems (GenAI, Agentic AI, traditional AI/ML systems, etc.)? If so, has the firm evidenced that it accounts for relevant differences in its approach depending on the specific functionality, risk profile, and governance needs of the AI system being used? Is the added value of complex systems, in terms of their enhanced capabilities, compared to simpler AI systems, analyzed? • Describe the assessment process and criteria used by the firm for determining which type of AI method is suited for a specific use case. • Before implementing new AI systems, what steps are carried out by the firm to ensure they are ready for deployment, including the principal regulator's consultation (where applicable)? • Describe how the AI system lifecycle is managed, including use of training data, monitoring performance, validation, testing and risk management. • What are the current AI use cases? How are unique risks related to specific use cases (e.g., hallucinations, bias) recorded and assessed?	• Specific risk management frameworks and processes for advanced AI systems (including model validation, testing, and monitoring; guardrails and harnesses; training, etc.).

Area	Considerations	Examples of Questions	Information Sources

6. AI model validation, testing and monitoring	• Does the firm conduct relevant evaluation and testing that is proportionate to the assessed risk materiality of the AI use case, system or mode?
	• Does the firm identify and assess evolving AI-specific risks such as model drift, adversarial attacks, data poisoning or new regulatory requirements?
		• How are AI-specific risks (including data quality, data management, algorithmic bias, model explainability, and operational resilience) systematically identified, assessed, monitored, and managed within the firm's risk management framework?
		  ○ What is the frequency of testing and oversight over the AI model? What reporting is provided regarding the AI model?
		  ○ Have you integrated preventative and post-event measures against hallucination risk?
		  ○ Have you determined the extent of transparency and explainability required of an AI use case, system or model according to its assessed risk materiality, and established the relevant controls accordingly?
		  ○ Have you defined what the firm considers "fair" outcomes and have appropriate controls to identify and mitigate harmful biases and discriminatory outcomes across the AI lifecycle, calibrated to its assessed risk materiality?
		  ○ Are key performance indicators of the AI model implemented and monitored?
		• How is efficient collaboration ensured between all relevant stakeholders in the AI lifecycle, i.e., reviews from model engineers, system operators, data scientists, business experts and senior management?
		• Is AI also deployed to govern or monitor AI products and, if so, how? How do firms evaluate the value of such use cases and how do firms retain control over the process?
			• Validation, testing, and monitoring exercise results of AI systems (incl. resilience and red team testing)
			• Monitoring dashboards (performance/fairness/drift), alert thresholds, rollback plans, usage and change logs.

7. Controls and human oversight of AI systems	• Does the firm have policies and procedures in place to define and differentiate roles and responsibilities for human-AI configurations and oversight of AI systems beyond	• How is human oversight implemented for AI systems and how does the firm evidence its effectiveness? For Agentic AI, also consider: • What are the specific risks associated with the data and systems the agent has access to, scope of actions the agent can take, the potential for the agent to share information and to whom, the reversibility of those actions, and the agent's level of autonomy?	• Second and third line of defense reviews (incl. Legal and Compliance) of AI systems (pre- and post-deployment) • External reviews of AI systems (e.g.,

Area	Considerations	Examples of Questions	Information Sources
	senior management?	• What are the specific controls for agents? How does the firm limit the scope of impact of agents (e.g., by designing appropriate boundaries at the planning stage)? • Is there sufficient clarity in the roles and responsibilities of both agents and humans throughout the overall lifecycle? Are there adaptations to appropriate human oversight to address automation bias associated with increasingly capable agents? • What new / additional technical controls and processes are in place for agentic components (e.g., pre-deployment and post-deployment testing, gradual rollouts, continuous monitoring? • How are end users informed of the agent's range of actions, access to data, and the user's own responsibilities?	system of controls reviews, audit controls, data controls)
8. Training and AI literacy	• Does the firm deliver role-specific training across relevant stakeholders (e.g., Board, senior management developers, validators, front-office, compliance) and promote AI literacy?	• How does the firm ensure it has the necessary knowledge, staff training/capacity, resources and compliance measures in place for AI systems, particularly regarding privacy and data security? • How does the firm ensure relevant staff can explain, challenge and escalate AI outcomes in their domain? • How does the firm assess and document the qualifications, expertise and ongoing training of individuals responsible for AI policy development, implementation, controls and monitoring to ensure they remain competent as AI technologies — and regulatory expectations — evolve?	• Mapping of AI related tasks with AI system owners, validators, and product owners, with training plans and competency assessments. • Training material, attendance, certifications and competency assessments for employees involved in all stages of the AI-system's lifecycle.

3.2 Third-party and outsourcing risk management

The importance of managing third-party risks as part of a robust governance framework is well-established, with accepted principles relating to outsourcing, operational resilience, and business continuity. As a general starting point, a supervised firm retains full responsibility, legal liability, and accountability to the regulator for all tasks that it may outsource to a service provider to the same extent as if the service were provided in-house. The following from the IOSCO Final Report on Principles on Outsourcing bears reiterating:[25]

- **Principle 5 (Concentration of outsourcing arrangements):** A regulated entity should be aware of the risks posed, and should manage them effectively, where it is dependent on a single service provider for material or critical outsourced tasks or where it is aware that one service provider provides material or critical outsourcing services to multiple regulated entities including itself.

- **Principle 6 (Access to data, premises, personnel and associated rights of inspection):** A regulated entity should take appropriate steps to ensure that its regulator, its auditors, and itself are able to obtain promptly, upon request, information concerning outsourced tasks that is relevant to contractual compliance and/or regulatory oversight including, as necessary, access to the data, IT systems, premises and personnel of service providers relating to the outsourced tasks.

While these foundational principles remain relevant, further practical tools can help address the distinct risk dimensions that AI systems introduce. In the more specific context of AI, IOSCO highlighted in the 2021 AI Report that regulators should promote practices so that compliance and risk management personnel in financial institutions understand and challenge the algorithms that are produced and conduct due diligence on any third-party service provider. The 2025 AI Report notes that the concentration of AI technology service providers presents specific risks, especially if advancements in AI are used in relation to algorithmic trading, robo-advising, and asset management.[26] At a firm level, concentration risk arises where the firm has not diversified its third-party providers and so may introduce single points of failure. More broadly, concentration risks may arise where multiple

[25] IOSCO Final Report on Principles on Outsourcing. Reference may also be had to FSB's Toolkit on Enhancing Third-Party Risk Management and Oversight.

[26] To note that this would be true for other material use cases, as the uses of third-party AI services evolves.

supervised firms use common service providers and so operational risks are correspondingly concentrated and may even increase to the extent that they present a systemic risk.[27] Detection and monitoring of concentration, outsourcing, and dependencies continue to present significant challenges for both financial institutions and supervisors alike.

Supervisory toolkit

Table 4 outlines various areas for consideration by supervisory authorities, including those highlighted in the Survey, in the oversight of third-party outsourcing risks in relation to the use of AI systems in financial products and services, including key considerations and examples of questions that supervisors may find useful.

[27] Examples of operational risks include: If the service provider suddenly and unexpectedly becomes unable to perform services that are material or critical to the business of a significant number of supervised entities, each entity will be similarly disabled; A latent flaw in the design of a product or service that multiple supervised entities rely upon may affect all these users; If multiple supervised entities depend upon the same provider of business continuity services (e.g., a common disaster recovery site), a disruption that affects a large number of those entities may reduce the capacity of the business continuity service.

Table 4: Evaluating third-party and outsourcing risks

Area	Considerations	Examples of Questions	Information Sources
1. Assessment of Risk-Proportionate Controls	• Does the firm have policies and procedures around onboarding, and development and deployment controls for third-party AI use, and are they adequate for the risk materiality of the use case, system or model? • Do policies and procedures cover whether testing of third-party AI products and services uses the firm's own data and use cases and do processes exist to receive notifications and assess the impact of third-party AI updates or changes? • Does the firm undertake appropriate compensatory testing (i.e., additional testing to address gaps, weaknesses or limitations in testing or disclosures from the third party)?	• How does the firm select Third-Party AI service providers? What due diligence does it undertake? • Does the firm keep a register of third-party contracts and usage focusing on AI-related providers? • Demonstrate how the firm's control framework scales with the risk materiality of different third-party AI applications. • What testing has the firm conducted using actual data, and how does the firm address informational gaps when providers don't disclose sufficient technical details? • Explain in detail the firm's process for evaluating and managing updates from third-party AI providers.	• Contractual terms and conditions relating to third-party vendors and their AI use. • Risk materiality classifications for all third-party AI use cases with corresponding control frameworks. • Evidence of testing protocols using the firm's own data and specific use case scenarios. • Documentation of compensatory testing addressing gaps from inadequate third-party disclosures. • Formal review schedules and update notification processes for third-party AI systems, presence of received notifications or checks in line with the firm's process. • Results of AI usage surveys, third-party registers and incident reports.

Area	Considerations	Examples of Questions	Information Sources
2. Transparency Measures	• Does the firm assess the level of transparency it receives from third-party AI providers on how key risks (data, model, technology, and cybersecurity) are addressed during development and deployment? • Where third-party transparency measures are not sufficient, has the firm employed compensatory measures such as additional testing, greater human oversight, or appropriate disclosures to users?	• How does the firm assess whether a third-party provider's transparency level meets established expectations? • What specific compensatory measures has the firm implemented where third-party AI lacks required transparency, and how does the firm validate their effectiveness?	• Evidence of regular review/updates to transparency expectations for third-party AI providers. • Records of transparency assessments conducted on third-party AI providers, documenting evaluation of their disclosures on data governance, system development, technology infrastructure, and cybersecurity controls. • Documentation of compensatory measures including additional testing, enhanced human oversight, and user disclosures.
3. Fairness Due Diligence Assessment	• Does the firm exercise due diligence regarding the fairness practices of third-party AI providers, recognising that the firm remains accountable for the fairness outcomes of third-party AI used within the organisation?	• How does the firm verify that third-party AI providers have appropriate risk management practices in place? • How does the firm conduct due diligence on third-party providers' fairness practices, and what evidence does it require? • Given that the firm remains accountable for fairness outcomes, how does the firm monitor and validate that third-party AI produces fair results across different client segments?	• Due diligence documentation on third-party AI providers' fairness practices and methodologies. • Evidence of ongoing monitoring for discriminatory outcomes from third-party AI systems. • Accountability frameworks demonstrating how the firm maintains responsibility for fairness of the outcomes.

Area	Considerations	Examples of Questions	Information Sources
4. Supply Chain Risk Assessment	• Does the firm ensure that key third-party and open-source AI models, datasets, and dependencies have undergone supply chain risk assessments and validation, including reviews of model provenance, training data integrity, and other known vulnerabilities?	• Explain in detail the firm's supply chain risk assessment for your most critical third-party AI systems, including model provenance and training data integrity. • How does the firm validate and monitor known vulnerabilities in its third-party and open-source AI dependencies?	• Third-Party risk management policy requirements in relation to AI related risks and subcontracting. • Documentation of supply chain risk assessments for key AI models, datasets and dependencies. • Documentation of model provenance reviews, training data integrity checks, and vulnerability assessments. • Evidence of validation processes (including testing processes) for both third-party and open-source AI components.
5. Concentration Risk	• Does the firm assess potential concentration risks arising from over-reliance on key third-party AI providers? • Does this assessment consider both direct dependencies (e.g., a single AI model or vendor) and indirect dependencies (e.g., reliance on shared infrastructure or cloud providers)?	• What analysis have you conducted on concentration risks from the firm's third-party AI providers, including indirect dependencies? • What specific steps is the firm taking to reduce over-reliance on key providers, and what are the target timelines?	• Analysis of direct and indirect concentration risks from key third-party AI providers. • Diversification strategies and implementation roadmaps. • Regular concentration risk monitoring and internal and regulatory reporting mechanisms.

Area	Considerations	Examples of Questions	Information Sources
6. Contingency Planning	• Does the firm have robust contingency plans to address potential failures, unexpected behaviour of third-party AI, or discontinuing of support by vendors, particularly for third-party AI used in high-risk materiality use cases, systems or models?	• If the firm's primary AI vendor discontinued support with minimal notice, what specific steps would the firm take and how long would full functionality restoration take? • Provide evidence of the firm's contingency plan testing for high-risk third-party AI applications. • What backup systems or alternative providers does the firm have validated and ready for deployment? • How does the firm assess potential negative consequences in case of discontinuity, i.e., data loss/alteration/leakage?	• Robust contingency plans specifically addressing third-party AI failures and unexpected behavior, including exit plans for normal and stressed exit scenarios. • Vendor discontinuation and vendor-switching response procedures with defined recovery timeframes. • Evidence of contingency plan testing, particularly for high-risk materiality use cases.
7. Legal Framework and Accountability	• Does the firm have legal agreements to facilitate clear expectations and responsibilities, including clauses pertaining to performance guarantees, data protection, the right to audit, notification when AI is introduced, or seeking the firm's agreement before incorporating AI?	• What contractual rights does the firm have to audit third-party AI providers, and how frequently does the firm exercise these rights? • How do the firm's legal agreements provide for notification and approval before AI is introduced or modified in products and services it uses?	• Updated legal agreements with clear performance guarantees and data protection clauses. • Contractual provisions for audit rights, AI introduction notifications, and approval processes. • Documentation showing how agreements facilitate clear expectations and responsibilities.

Area	Considerations	Examples of Questions	Information Sources
8. Capability Development	• Does the firm build awareness and develop capabilities of staff involved in procurement, development, deployment, and use of third-party AI systems?	• What specific capabilities has the firm developed in the firm's procurement, risk, and business teams for managing third-party AI systems (or products)? • How does the firm ensure staff awareness and competence keeps pace with your expanding use of third-party AI systems?	• Training programs and competency frameworks for staff involved in third-party AI procurement and management. • Evidence of appropriately timely and updated awareness-building initiatives across relevant business functions. • Capability assessment results and development plans.
9. Use of Complex AI Products	• Does the firm conduct enhanced assessments when using more complex or novel third-party AI products and services that the firm may have less experience with?	• For the firm's most complex third-party AI implementations, explain in detail its enhanced assessment process and additional safeguards. • How does the firm determine when a third-party AI product requires more detailed assessment, and what additional measures does the firm apply?	• Enhanced assessment protocols for complex third-party AI products and services. • Documentation of additional due diligence measures for unfamiliar or sophisticated AI systems (or products). • Evidence of expert involvement and extended evaluation periods for complex implementations.

3.3 Disclosure

As AI systems become increasingly embedded in financial services and material to the operations of financial institutions, supervisors will need to assess whether investors and other stakeholders are receiving appropriate disclosures regarding these AI systems, their use, and the associated risks. Clear and timely disclosure enables investors to better understand how AI systems are integrated into financial products and services and can be a critical tool for promoting transparency and integrity in capital markets. This transparency supports informed decision-making by investors and helps maintain confidence in the integrity of markets.

Disclosure also plays an essential role in assisting investors and clients with identifying and assessing risks associated with AI systems. These risks may include data quality concerns, algorithmic bias, hallucination, and potential broader impacts on market fairness and stability. Through reviewing appropriate disclosures, investors and clients are better positioned to consider such risks when making an investment decision or choosing to engage with a particular service or product.

At the same time, disclosures should be calibrated to avoid revealing information that could compromise cybersecurity, intellectual property, or the resilience of critical systems. While high-level, outcome-focused disclosures may be appropriate for investors, detailed technical information about AI system architectures, model vulnerabilities, or proprietary methods may introduce operational or security risks if made public and may not be of use to investors. Therefore, supervisory oversight of disclosures should strike an appropriate balance: providing meaningful, comprehensible information that enables stakeholders to understand the role of AI in financial products and services, while avoiding the release of sensitive information that could heighten security risks or undermine firms' competitive positions. What constitutes appropriate disclosure will vary by regulatory framework, but supervisors may want to consider the materiality of information disclosed to the investor, technical or commercial sensitivity of information, security risks associated with disclosure, and potential impact on market confidence.

Challenges to comprehensive disclosure practices include: disclosures miscalibrated for the audience; inconsistent terminology (such as lack of clarity on which systems are categorized as AI systems and which are not); lack of clarity on where and how AI systems are deployed; and insufficient disclosure of third-party AI components (where firms rely on third-party providers to provide disclosure when lacking internal AI expertise). Additional challenges arise from the use of generic or promotional statements; insufficient disclosure of material risks; the use of generic or boilerplate language that lacks firm-specific information that accurately reflects how AI is used and the associated risks; and imbalanced presentation of benefits and risks relating to the use of AI systems.

Box 4: Identifying misleading, inaccurate, or false claims relating to Artificial Intelligence

Firms can potentially make misleading, inaccurate, or false claims relating to the use of AI in capital markets. These claims can relate to, e.g., AI development, implementation, use, functions, and performance. These claims can appear in, e.g., filings, reports, websites, and social media. This conduct is sometimes referred to as "AI-washing," and it can be the basis for enforcement actions against market participants in certain jurisdictions. The Survey revealed that respondents have observed instances of supervised firms failing to adequately disclose the use of AI systems in the management of investment portfolios, making misrepresentations about AI-related investment activities, making false statements as to the level of implementation of AI systems in investment research or selection of securities, and reporting false performance attributed to AI systems. This conduct has been observed by respondents, specifically with respect to certain asset managers, fund managers, and issuers.

Example of incomplete disclosure leading to AI-washing[28]

The following is an example of AI-washing, i.e., misleading disclosure that was included in a reporting issuer's continuous disclosure documents.

"The company utilizes the most advanced AI technology. The company's warehouse houses the most sophisticated AI robotics. The company uses AI to solve world issues. The company's use of AI modernizes the company's business processes and will disrupt the industry in which it operates. The company's business operates in a leading global artificial intelligence domain. The company in its public filings only discussed its acquisition and development of AI technology and it appears to be the only business of the issuer."

The issues with these statements are as follows. First, in the above example, the issuer made potentially misleading and overly promotional claims regarding the capabilities of its technologies, without being supported by facts and corporate activities.

Second, the issuer described itself as being a global leader and disruptor despite having generated only nominal revenue from its operating activities. Making broad statements without supportable financial statement performance measures and additional detail regarding the specific aspects of

[28] Adapted from CSA Staff Notice 51-365 Continuous Disclosure Review Program Activities for the Fiscal Years Ended March 31, 2024 and March 31, 2023.

its business or how the capabilities of the business will be measured and evaluated is misleading and promotional.

Third, the company's continuous disclosure record focuses entirely on AI technology; however, on review, substantially all the company's revenue came from the sale of general appliances. This is misleading because an investor reading the company's continuous disclosure record could reasonably assume that all the company's revenue is from its AI activity.

Supervisory toolkit

Table 5 outlines various areas for consideration by supervisory authorities, including those highlighted in the Survey, in the oversight of the disclosure to investors of the use of AI systems in financial products and services, including key considerations and examples of questions that supervisors may find useful.

Table 5: Evaluating disclosure of AI use

Area	Considerations	Examples of Questions	Information Sources
1. Disclosure of AI use in products and services to end-users	• Does the firm transparently inform clients when they are interacting with AI systems (e.g., chatbots, robo-advisors, or decision support tools)? • Does the disclosure clearly explain how AI systems work, including material information about their performance, limitations, and suitability?	• How does the firm ensure clients know when AI is generating responses or analysis? • Provide examples of how outputs are marked as AI-generated or synthetic. • How does the firm communicate AI decision logic and the firm's application of outputs in client communications or disclosures? • What internal controls exist to validate claims about AI sophistication, performance, or level of integration before they appear in disclosure documents? • How does the firm disclose material AI-related risks (e.g., hallucination risk, third-party dependencies, bias, cybersecurity incidents)? • How often is disclosure occurring (at a single point in time vs. whenever a client interacts with an AI system)? • How does the firm ensure that disclosure about AI use is accurate, evidence-based, and not promotional in tone (i.e., does not constitute AI-washing)?	• Client-facing materials (websites, apps, account agreements) stating whether AI was involved in the process, evidence of machine-readable labels or markings on AI-generated content, and consent language in privacy policies or service agreements specifying AI use. • Substantiation records for claims made in disclosures.
2. AI governance and strategy disclosures	• Does the firm publish disclosures on board/senior management oversight, governance and risk management frameworks,	• Does the firm publicly describe the Board's/Senior Management's role in overseeing AI systems? • Which senior managers are accountable for AI governance and risk management, as stated in the firm's public disclosures?	• Public reports documents such as annual reports or prospectuses describing use of AI in operations, internal controls, risks, and public incident reports; regulatory breach disclosures.

Area	Considerations	Examples of Questions	Information Sources
	and roles and responsibilities for AI?	• What internal review process does the firm use to validate marketing, investor, or governance disclosures relating to AI performance, system sophistication, or automation levels?	• Client facing documents and disclosures regarding the oversight of AI use by the firm. • Board/senior management reporting on risk disclosures referencing AI-enabled systems, accuracy issues, bias management, or fraud, summaries of data, architecture, performance, limitations, and bias metrics for each AI system.
3. Disclosure and consent	• Where AI systems, whether proprietary or run by a third-party, impact client rights, account access, or privacy, is the firm informing clients and seeking explicit consent or allow opting out?	• How does the firm inform clients and obtain consent when AI affects their products, the investment objective or strategy, pricing, or data rights or privacy? Where applicable, how does the firm obtain client consent? • Provide examples of where the firm offers explicit opt-in or opt-out for AI-driven decisions. • What contest or redress mechanisms does the firm have in place for clients to be able to challenge decisions/outcomes where AI has been used? • Where there is material reliance on third-party tools, do firms disclose such dependencies, including associated risks, rather than deferring transparency to technology vendors?	• Records evidencing opt-in/out mechanisms from AI services/products for clients potentially impacted by AI systems. • Client facing documents and disclosures regarding the use of AI systems, including client data used, use of AI systems on services provided to the client.
4. Identifying misleading claims	• Are AI related claims materially inaccurate, false, vague, or misleading, whether in filings, reports,	• How does the firm ensure that public statements about your AI are accurate, evidence-based, and not overstated?	• Board/senior management reporting on risk disclosures referencing AI-enabled systems, accuracy issues, bias management, or fraud, summaries of data, architecture, performance,

Area	Considerations	Examples of Questions	Information Sources
	marketing materials, websites, and social media? • Is the firm able to fairly and accurately substantiate its AI-related claims?	• Does the firm benchmark your AI systems' performance when making public claims? If so, what evidence underpins such claims? • How does the firm ensure that statements about proprietary AI tools accurately reflect the extent of third-party or vendor reliance? • How does the firm monitor and correct outdated public descriptions of the firm's AI use (e.g., if a system is scaled back, retired, or replaced)? • What governance mechanisms exist to detect and address potential AI-washing identified by clients, investors, employees, or external parties?	limitations, and bias metrics for each AI system. • Policies and procedures to review the use of AI-related claims in marketing materials, client documents, public statements; include documentation evidencing review of AI claims or statements.

3.4 Recordkeeping and reporting

Typically, most capital markets' regulatory frameworks require market participants to keep records of important daily decisions to support auditability, transparency, and explainability of those decisions. The use of AI systems, particularly Agentic and Gen AI systems, by market participants challenges these traditional requirements as – in some cases – AI-generated outputs may be complex and difficult to explain.

To promote auditability, transparency, and explainability, regulators often require market participants to document the use of their AI systems. Adoption of AI systems should not — as far as possible — avoid recordkeeping requirements. Documentation and logs should be maintained throughout the different lifecycle stages: design, development, modification, deployment, ongoing monitoring and retirement. Such documentation and logs should facilitate the effective supervision and monitoring of the use of AI systems and support market participants' accountability for the AI systems' output, proper functioning throughout the AI system lifecycle, and the implementation of corrective actions when adverse outcomes occur. Any such negative outcomes should be recorded and root cause analysis performed to determine their cause and any remediation required. Supervisors may require expert specialists and additional technical tools to review these records and documentation pertaining to AI systems.

Reporting refers to how supervised firms communicate information on their supervised activities to supervisory authorities. Comprehensive and regular reporting can assist supervisors in maintaining appropriate oversight of activities that may give rise to risks to the market and its participants and to their statutory objectives. As AI systems become embedded in core functions – such as trading, risk management, client communications and onboarding, trade surveillance, anti-money laundering and operational processes – reporting frameworks enable supervisors to capture information relevant to the use of AI.

Effective and periodic reporting by institutions could also assist supervisors in identifying emerging trends, risks, and systemic vulnerabilities stemming from the use of AI systems, which could adversely impact investor protection, market integrity, and financial stability. This can include visibility over the scale and materiality of AI deployment, reliance on third-party providers, the controls governing data and system inputs and outputs, and how systems are tested and monitored in production environments. Timely reporting, with the appropriate level of detail, also supports early identification of incidents, system failures, or data integrity issues that could affect clients, markets, or the institution's operational resilience.

As the use of AI continues to grow, supervisory authorities may consider reviewing whether existing reporting requirements remain fit for purpose given challenges that may stem from emerging AI technologies and techniques. Such a review could involve enhancing existing regulatory requirements and expectations, establishing dedicated or revised AI reporting templates, or integrating AI-related data points into broader operational or technology risk reporting frameworks. Whatever the approach, reporting should aim to provide supervisors with sufficient insight to exercise effective oversight while avoiding unnecessary burden on firms.

Supervisory toolkit

Table 6 outlines various areas for consideration by supervisory authorities, including those highlighted in the Survey, in the oversight of the recordkeeping and reporting, where appropriate, of the use of AI systems in financial products and services, including key considerations and examples of questions that supervisors may find useful.

Table 6: Evaluating the recordkeeping and reporting of AI use

Area	Considerations	Examples of Questions	Information Sources
1. Documentation of AI Lifecycle Oversight	• Does the firm maintain comprehensive documentation covering all stages of an AI system's lifecycle, including, e.g., design and planning; data collection and processing; architecture/model choice, model building, training, testing and validation; deployment, operation, modification and monitoring; algorithms used; and decommissioning plan?	• What mechanisms ensure comprehensive traceability and auditable recordkeeping for AI-driven decisions? • Does the firm use different tools for traceability and audit depending on the AI systems type (e.g., ML or LLM based)? • How does the firm record and track modifications to AI systems over time? Does the firm maintain a copy of the models used and modified during the time (model versioning)? • What processes ensure lifecycle documentation remains current and complete both for proprietary AI systems and third-party AI systems?	• Documentation and logs across the AI system lifecycle. • Records of AI model choices per use case, or for proprietary models, model architecture choices and rationale. • Change management records for AI system updates.
2. Documentation of Business Objectives Related to AI	• Does the firm clearly document the business objectives behind AI adoption, including expected benefits, risk considerations, and alignment with strategic goals?	• What business objectives does this AI system support, and how are they documented? • How does the firm ensure AI objectives align with your overall risk management framework?	• Business case documents for AI projects, including any board-approved papers referencing AI objectives.
3. Documentation of AI-Generated Outcomes	• Does the firm maintain records of AI-generated outcomes, focusing on investor or market impacting use cases?	• What processes are taken to review and document AI outputs on a periodic basis? • What audit trail does the firm maintain for AI-system outcomes in investor or market impacting use cases?	• Documentation and logs of validation, testing, and monitoring procedures, internal reports, and documentation of

Area	Considerations	Examples of Questions	Information Sources
		• What processes ensure traceability between AI outputs and final actions taken? • How does the firm document and retain evidence of AI-generated outcomes? • How does the firm document and retain evidence of human oversight or intervention in decisions made by AI systems, particularly in investor or market impacting AI applications?	algorithms and decision pathways used in AI systems.
4. Documentation of Explainability of AI Outputs	• Does the firm maintain records that explain how AI systems generate outputs, including decision-making logic and reasoning?	• To what extent can the firm meaningfully explain the outcomes of the AI systems? How are these explanations tailored for different stakeholders (end-users, Board and senior management, supervisors)? • How does the firm ensure explanations are accurate, fair, and free from bias taking into account the explainability techniques used to enhance human understanding of the Ai systems? • How does the firm document the logic behind AI-generated decisions? • What records demonstrate the firm's ability to explain AI outputs to stakeholders? • What evidence does the firm maintain to show that explanations of AI outputs are accessible and understandable to non-technical stakeholders, such as retail investors? • Does the firm use different tools for tracing explainability and interpretability?	• Internal guidelines for interpreting AI outputs and audit trails showing human validation of AI outputs.
5. Records on Incidents relating	• Does the firm have policies and procedures for recording and reporting incidents that note	• Are incident logs and monitoring records maintained in a manner that enables effective supervisory review and supports accountability for AI system outcomes?	• Policy for recording and reporting AI-related incidents.

Area	Considerations	Examples of Questions	Information Sources
to AI products or services	whether these incidents included the use of AI systems? • Does the firm keep records of operational failures, misconduct, or security breaches, and maintain detailed records of AI-related incidents, including root cause analysis and remediation actions?	• Does the firm periodically review and update its incident management policies and procedures to reflect lessons learned from past AI-related incidents and evolving regulatory expectations? • What thresholds does the firm maintain for an AI-related incident to be reportable? • Does the firm's business continuity procedures consider a response plan to incidents (e.g., identification of an incident, steps taken to address and remediate, any escalation procedures or steps, documentation of an incident)?	• Records of AI-related incidents over the past 12 months. • Records of remediation.
6. Reporting	• Does the firm comply with relevant jurisdictional reporting requirements, to provide the supervisory authority with appropriately frequent, timely and high-quality reporting across the areas listed above?	• What reporting processes are in place to ensure timely and high-quality reporting on AI in line with relevant reporting requirements? • What constitutes an AI-related reportable incident within the firm? Does the firm define this in its internal rules or guidelines? • What reporting mechanisms cover third-party or external data dependencies? • How does the firm notify supervisors of significant AI failures, biases, cybersecurity events, data breaches, or incorrect outputs? • What criteria determine whether an AI system requires mandatory notification or prior approval?	• Incident management policies, including those specific to AI. • Templates for incident reporting to authorities. • Incident documents and logs with timestamps and impact assessments, including root cause analysis reports for AI failures.

3.5 Monitoring AI use

As AI systems are increasingly deployed in capital markets, the identification and development of consistent indicators of firms' use of AI becomes an increasingly important part of the supervisory process. Such indicators support a risk-based and proportionate approach to regulation and supervision by enabling authorities to better understand patterns of AI adoption, identify areas of heightened risk, and prioritize supervisory engagement. The monitoring of indicators enhances supervisory oversight of the use of AI and enables early identification of activities that may undermine investor and market protections.

Table 7 outlines examples of key indicators that supervisors may find useful to refer to in their oversight of AI use in financial products and services. These include indicators highlighted through the results of the Survey as being used by responding IOSCO members.

Table 7: Key indicators for supervisory oversight of AI use

Indicator	Purpose
AI adoption	
Proportion of supervised firms that use AI technologies	This provides an overview to the supervisory authority of how widespread AI use is across supervised firms and use cases. Through understanding AI use across sectors of firms, supervisors can analyse AI adoption patterns.
Proportion of use cases in development vs those deployed in production	This distinguishes between experimental projects and active systems affecting business outcomes, to assess current vs. emerging risk exposure. Supervisors could classify firm usage in three categories: (i) using AI in production, (ii) experimenting with AI and (iii) not using AI, to facilitate monitoring of: • Conversion rate of experimentation to production deployments. • Growth rate of production use cases. • Adoption rate of AI based on firms moving from not using AI to experimenting with AI to using AI in production.
AI technologies Research and Development (R&D) spending	By understanding the amount invested in AI technologies R&D over past years or forecast for future years, supervisors can understand the likely growth trajectory in AI use in supervised firms, as well as its relative priority to a particular firm.

Indicator	Purpose
Number of AI-related patents/patent applications by a firm or sector	Similar to R&D spending, this metric can provide an indication of internal research and investment focused on AI and whether proprietary AI systems are being prioritized by a particular firm or sector.
Nature and use cases of AI systems	
Detailed inventory of the AI use cases and AI system type in development and deployment	By categorizing AI systems by use case and system type, supervisors can understand how AI is being used across firms in a spectrum of use cases.
AI system-related attributes	Information concerning core attributes of an AI system, including specific underlying models, technical specifications, and linkages to products/services can promote an understanding of how an AI system is used and its potential risks.
Third party dependencies and concentrations	Supervisors may also wish to consider information on the proportion of AI systems and use cases of those systems that are supported through third party AI service providers vs internally developed. Supervisors may also want to understand across firms the level of dependency on and concentration in specific third-party providers for specific AI systems, as well as the dependencies on and concentrations in those third-party providers for the firm as a whole (e.g., what other systems or services the third-party provides to the firm).
AI system performance	
Frequency and severity of incidents in AI systems	Information tracking failures, errors, outages, cyber incidents, data breaches, instances of frauds/misconduct and malfunctions in AI systems can identify reliability and safety concerns, including their business impact, duration, and affected investors/end users.
Model performance validation and monitoring metrics	Information concerning a firm's ability to identify and track key performance indicators such as prediction accuracy, error rates, and model drift and degradation over time can enable early identification of underperforming systems.
Examples of sector specific metrics[29]	

[29] These are examples of indicators that may help monitor specific risks of AI use cases in each sector from the 2025 AI Report but are not exhaustive, nor are they intended to be prescriptive.

Indicator	Purpose
Broker-dealers	Incidents or alerts of AI-driven mis-selling or unsuitable recommendations requiring human intervention. Bias or fairness indicators in recommendation or communication tools. Input/output correlation metrics, checking whether small changes in input lead to disproportionately large, erratic, or unexplained changes in system output.
Asset Managers	Accuracy or consistency of AI-supported valuations relative to market benchmarks. Incidents or alerts of AI-driven mis-selling or unsuitable recommendations. Indicators of mandate or portfolio-construction drift attributable to AI-generated outputs. Level and frequency of human intervention in AI-driven investment process. Accuracy or consistency of risk profiling of clients, and suitability of AI-supported asset allocation and investment recommendation relative to client profiles in the context of robo-advisers.
Exchanges and Other Financial Market Intermediaries	Market quality measures (e.g., bid-ask spreads, price impact) where AI-driven systems influence order routing or matching. Effectiveness of AI-driven surveillance systems in detecting market manipulation, insider trading, and other misconduct. Market volatility or instability patterns associated with AI-driven trading systems. Accuracy or consistency of AI-supported models for financial risk management by market infrastructure providers. Trade matching and settlement failures arising from AI automation in processes.

Data sources

To support oversight of financial institutions and collection and analysis of the indicators highlighted above, supervisors may consider a variety of data sources,

tools, and techniques. As highlighted in the FSB 2025 report[30] and corroborated by the Survey responses, data is most often collected through:

- Supervisory reporting and data collection during examination and inspection: Data collected through supervisory reporting, examination, and inspection can include AI-related data where feasible. Such data, especially when collected on a regular basis, can enable quantitative and qualitative analysis of AI use by supervised firms.

- Regular and structured outreach and engagement with firms through supervisory engagement: Direct engagement with firms can enhance understanding of how AI systems are being used in financial products and services and where risks may emerge, while at the same time reinforcing regulatory requirements and expectations.

- Surveys of market participants: While surveys are used infrequently by supervisory authorities, according to the Survey results, sector-wide surveys can promote effective monitoring, covering AI inventories, third-party dependencies and concentrations, and AI-related investments, with survey designs that remain consistent yet adaptable, timely, and forward-looking to capture emerging developments.

- Research on publicly available information: Public reports (e.g., annual reports, press releases, company disclosures, industry, and sector reports) can provide information ahead of more focused engagement.

- Real-time monitoring tools used by supervisors: Only a few authorities reported employing real-time capabilities, including systems to analyze trading behavior through AI-powered detection tools. Real-time monitoring implies a cost-benefit assessment, but cross-border information sharing can help reduce the costs and frictions in developing such capabilities.

- AI tools: While most IOSCO members responding to the Survey indicated that they are not currently using AI systems to support their monitoring activities, a few do use AI systems for analyzing data from routine monitoring or public filings, to detect outliers, anomalies, or to scan available data for relevant information. Analyses reportedly are carried out by using both commercial and internally developed tools.

[30] FSB 2025 report.

- Enhanced information sharing among national or cross-border authorities**:** Such information sharing can help to minimize duplication and improve coverage and comparability.

- Sandboxes, innovation hubs, and other initiatives: Some IOSCO members are using other ways to gather information on AI use in firms, including through engagement and outreach to industry, market participants, and others.[31]

[31] For instance, the UK FCA has launched its AI Lab initiative and live AI testing service: FCA set to launch live AI testing service | FCA.

Conclusion

AI systems have the potential to reshape capital markets, offering opportunities to improve efficiency, risk management, and service quality, while also introducing new and evolving risks. As AI technologies become more capable, more complex, and more deeply embedded in financial services – including through the increasing use of GenAI and early forms of Agentic AI – adequate supervisory oversight of the application of AI in capital markets remains a crucial way to mitigate those risks, while supporting the prudent realization of AI's potential benefits.

As technologies evolve, new risks may emerge – from hallucination and explainability challenges, to autonomy and control issues, to novel forms of bias, manipulation, or operational vulnerability. Supervisors will need to monitor developments closely and adjust their approaches accordingly. The principles of proportionality, accountability, transparency, and reliability remain central to maintaining trust in markets as AI adoption grows.

The supervisory toolkit presented in this report provides IOSCO members with practical tools to support proportionate, risk-based oversight of AI systems used by supervised firms. It is designed to complement, not replace, jurisdictional frameworks, and to offer a common foundation for supervisory dialogue between authorities and firms. Given the rapid evolution of AI systems and their use cases, it may be prudent to periodically review whether the toolkit remains fit for purpose.

IOSCO remains committed to supporting its members in navigating these developments. Through the continued work of the AIWG – including the forthcoming review of emerging industry practices on disclosure, recordkeeping, reporting, and governance – IOSCO will continue to support responsible innovation while promoting market integrity, investor protection, and financial stability.

IOSCO welcomes feedback from stakeholders on the supervisory toolkit set out in this report. Respondents are invited to share their views on the tools in the report as well as any emerging practices in industry. Feedback may be submitted via a short survey by 26 June. IOSCO will take responses into account as it continues to develop its work on AI, including in the next phase of its review of emerging industry practices.

Annex I: IOSCO Members' Survey Findings

This chapter presents more detailed results from the Survey. Each section focuses on key findings and draws out notable examples from responding IOSCO members. It covers:

- Further observations on AI-related technological developments
- Regulatory frameworks
- Governance
- Disclosure
- Reporting
- Recordkeeping
- Third-party risk
- Inspections/examinations
- International cooperation and private sector engagement

AI-related Technological Developments

In addition to the high-level AI-related technological developments identified in Section 2.1, the Survey outlined more detailed insights from IOSCO members that responded to the Survey, particularly on the drivers and benefits of AI adoption, and barriers to further adoption.

Key trends in AI-related technological developments

Surveyed members noted a trend towards increasing experimentation around GenAI and Agentic AI, citing the increased interest in these technologies as the most significant technological trends in financial services since the 2025 AI Report. These recent and emerging forms of AI are increasingly being explored by firms, through trial and testing. The Survey indicated that firms are broadly taking a cautious approach to AI adoption for GenAI and Agentic AI, particularly for uses that would incorporate such AI systems into core decision-making in financial products and services. Primarily, examples of GenAI adoption focused on back-office functions. However, surveyed members indicated a need for increased focus on governance and accountability for GenAI systems.

AI adoption

As noted in the 2025 AI Report, while some jurisdictions have reported that the use of AI in the financial sector is increasing, particularly for some analytics such as market analysis, forecasting and portfolio optimization, the deployment of

GenAI in the financial sector appears focused primarily on support and risk-related functions, rather than in high-stakes decision-making. The Survey results highlight relatively strong uptake in areas such as fraud detection and prevention, risk management, compliance, client service and client engagement.

One member reported that AI is commonly applied in algorithmic high-frequency trading (HFT), primarily for trading signal generation (including deep learning, reinforcement learning and GenAI), and that AI is not commonly used directly in trade order routing, execution, or cancellation.

Increased adoption in the financial sector is supported by other quantitative findings taken from other surveys and analyses. For example:

- 85% of 118 firms surveyed in the UK are using or planning to use AI (Bank of England & FCA, 2024).[32]

- 55% of 728 firms surveyed in the EU are using or planning to use AI. However, there is a pronounced disparity across firm sizes, with adoption increasing to 96% for large firms but dropping to 35% for micro firms (ESMA, 2026).[33]

- 55% of AI use cases in the UK involve some automated decision-making, though only 2% are fully autonomous (Bank of England & FCA, 2024).

- 33% of use cases rely on third-party solutions; 17% involve foundation models (LLMs) (Bank of England & FCA, 2024).

- GenAI adoption (use of GenAI in at least one use case) in surveyed financial institutions is expected to rise from 75% to 87% within 3–5 years in Hong Kong (HKIMR, 2025).[34]

- Growing AI adoption among 40 market intermediaries (ASIC, 2025).[35]

[32] Bank of England & Financial Conduct Authority. (2024). Artificial intelligence in UK financial services: Survey report. Bank of England. Retrieved from Artificial intelligence in UK financial services - 2024 | Bank of England.

[33] TRV Risk Monitor report and TRV Risk Analysis (February 2026): ESMA50-481369926-30599 TRV article - AI adoption and trends in securities markets: EU evidence.

[34] Hong Kong Institute for Monetary and Financial Research. (2025). Financial services in the era of generative AI: Facilitating responsible adoption. Hong Kong Monetary Authority. Retrieved from genairep1.pdf.

[35] Australian Securities and Investments Commission (2025): **MIU - Issue 166 - April 2025 | ASIC**.

Barriers to enhanced adoption of AI as identified in the Survey

Surveyed members highlighted the following barriers to increased adoption of AI:

- **Governance gaps**: Surveyed members reported that firms struggle to establish robust AI governance frameworks (see Chapter 3 for more detail and relevant supervisory tools).
- **Reliance on third-party models**: Some surveyed members report that given that smaller firms are more likely to depend on external vendors and cloud-based services, the costs, risks, and lack of understanding of these third-party dependencies can create a barrier
- **Talent shortages and competition for talent**: Surveyed members cited a lack of skilled professionals in AI and related fields as a key limitation to financial firms' ability to develop, implement, and maintain advanced AI systems.

Looking Ahead

The Survey results highlight that the adoption of GenAI in financial services is accelerating, particularly for back-office applications. Survey results indicate that financial institutions are preparing for a notable increase in the integration of GenAI across business services. Several respondents anticipated further expansion of GenAI into a wider range of business functions (including compliance support, risk analysis/modelling and client-facing processes), as organizations move beyond pilots toward more structured deployments; however, responses generally described these expectations qualitatively and with differing time horizons. This trend appears driven by the expectation that GenAI systems will enhance efficiency, improve risk detection, and streamline client interactions.

Surveyed members have expressed growing concerns over GenAI and Agentic AI systems potentially introducing new governance and oversight challenges for firms and regulators.

An on-going significant trend appears to be an increasing reliance by firms on third-party vendors, which could heighten outsourcing, concentration, and third-party dependency risks across the sector.

The Survey results emphasize the need to understand and continue to monitor AI-related developments in the financial sector to effectively address these potential risks.

Additionally, the Survey results highlight the importance of cross-border collaboration and information sharing among regulators to promote shared understanding of developments and to proactively manage risks that may transcend national boundaries. As AI continues to be explored and implemented in financial products and services, ongoing coordination and proactive

engagement among regulators will play an important role in supporting market integrity, investor protection, and financial stability.

Regulatory frameworks

The majority of surveyed IOSCO members have indicated that their respective regulatory frameworks cover the use of AI by financial participants. Results of the analysis from jurisdictions surveyed in IOSCO's 2025 AI Report indicated that approximately 75% of jurisdictions had adopted a regulatory approach over AI uses. Current results show a notable increase, where approximately 90% of respondents now indicate that they have a regulatory framework in place addressing uses of AI by financial participants.

Almost all respondent IOSCO members indicated they take a technology-neutral approach, reporting that existing sectoral regulations cover key areas such as governance, operational resilience, disclosures and reporting, regardless of the technology being used. Therefore, though jurisdictional frameworks may not explicitly mention AI uses, AI uses in financial products and services are often addressed via supervisory work, examinations, or the publication of guidance reminding supervised firms of their obligations under existing sectoral regulations.[36]

The majority of responses to the Survey (85%) indicate that the frameworks applicable to AI uses adopt a risk-based approach, and a minority of responses (35%) indicate that their framework follows a use-case specific approach, meaning that in such instances regulatory frameworks either cite or prohibit specific use cases. One further observation is that principles-based regulatory frameworks, as opposed to rule-based, like the EU AI Act, seem to not prescribe or prohibit specific use cases, though in practice certain uses may be inoperable due to a lack of meeting the required standards.[37]

[36] In May 2024, ESMA made a public statement on the use of AI in investment services, reminding firms of their obligations under existing EU sectoral regulations, such as MiFID. In September 2025, the FCA also published its AI update, indicating its approach to support the responsible adoption of AI by firms in the UK financial sector. In the US, the SEC staff published guidance relating to disclosure and other obligations, focusing on robo-advisors. See IM Guidance Update: robo-advisers, No. 2017-02, February 2017, https://www.sec.gov/investment/im-guidance-2017-02.pdf; Risk Alert, SEC Division of Examinations, Observations from Examinations of Advisers that Provide Electronic Investment Advice, November 9, 2021, https://www.sec.gov/files/exams-eia-risk-alert.pdf.

[37] Canada's OSC & QAMF indicate that certain activities, such as fully automated portfolio management services, may not meet required regulatory standards, due to a lack of accountability, explainability, or meeting investor protection requirements.

In the European Union, the EU AI Act, which concerns eight respondents,[38] adopts a risk-based approach[39] for uses of AI systems across the EU, where these are assessed based on the risks they pose to the life, health and fundamental rights of EU citizens. Though it applies to all sectors of industry (and thus not to financial services specifically), it identifies two particular use cases provided in the context of financial services which it categorizes as "high risk" in relation to the rights of EU citizens. Further detailed analysis of the application of the Act is provided in Box 4 below.

Outside the application of sectoral regulation for financial services participants, several respondents have also referred to other regulations or guidelines published by non-financial sector specific institutions that relate to the use of AI, such as national cyber-security or other regulatory bodies (i.e., the NIST AI Risk Management Framework).[40]

Regarding future initiatives, around half of the respondents to the Survey indicated that they are not planning to specifically regulate AI uses within financial services, though a number of jurisdictions (30%) also indicated that they plan to launch regulatory initiatives in this area. For instance:

- Saudi Arabia's CMA has indicated that it will consider its process for regulatory intervention once it has conducted a first evaluation of the use of AI in financial activities.
- Singapore's MAS recently issued a consultation paper proposing a set of Guidelines on AI Risk Management[41] to guide financial institutions on the responsible use of AI in the financial sector, setting out MAS' supervisory expectations on oversight of AI risk management in FIs, key AI risk management systems, policies and procedures, key AI life cycle controls, as well as capabilities and capacity needed for the use of AI.
- India's SEBI is in the process of formulating a framework of guiding principles, based on stakeholder feedback and inputs received following

[38] AMF (France), BaFin (Germany), HCMC (Greece), CBI (Ireland), CONSOB (Italy), CNMV (Spain), FI (Sweden), and AFM (The Netherlands).

[39] The risk-based approach under the EU AI Act considers risks to the fundamental rights of EU citizens.

[40] AI Risk Management Framework | NIST, see also the document published by France's ANSSI on building trust in AI through a cyber risk-based approach, and co-signed by a number of cyber-security agencies on the subject: https://messervices.cyber.gouv.fr/documents-guides/high_level_risks_analysis_ai_paris_summit.pdf.

[41] See publication of AI Risk Management Guidelines by MAS: https://www.mas.gov.sg/news/media-releases/2025/mas-guidelines-for-artificial-intelligence-risk-management.

its consultation paper on guidelines for responsible usage of AI and Machine Learning in the Indian securities market.[42]

Box 5: The EU AI Act

The EU AI Act is a legislation introduced into EU law in July 2024, which governs the use of AI systems in the EU, or those having an impact on the EU.

The EU AI Act aligns with the OECD definition of AI systems (discussed in Chapter 1) and applies it "horizontally" across all sectors of the EU market, including the financial sector, provided they use AI systems.

The EU AI Act's scope is broad, particularly concerning:

- **The actors involved**, as the legislation covers the entire AI value chain, including both providers and deployers of AI systems.

- **The technologies involved**, as the legislation covers all AI systems, providing for a definition closely aligned to that of the OECD[43], and targets certain so-called "general-purpose" AI models, which are capable of performing a wide range of distinct tasks and are trained using very large datasets.

Depending on the type of system and the risks to the fundamental rights of EU citizens associated with its use, the actors placing it on the market or using these technologies will be subject to certain obligations. Obligations also exist for providers of general-purpose AI models.

The EU AI Act classifies AI systems according to the risk levels their uses pose to citizens' life, health, and fundamental rights, and defines obligations based on these risks. The concept of "risk" in the Act therefore specifically targets risks that could harm individuals (manipulation, infringement of freedoms and privacy, social categorization, etc.) and not risks specific to the financial sector (market risk, financial stability, etc.). Depending on the risk level of the systems,

[42] See: https://www.sebi.gov.in/reports-and-statistics/reports/jun-2025/consultation-paper-on-guidelines-for-responsible-usage-of-ai-ml-in-indian-securities-markets_94687.html.

[43] Defined as "automated systems designed to operate with varying levels of autonomy and capable of adapting after deployment, and which, for explicit or implicit purposes, infer from the data they receive how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments."

the obligations arising from the Act are based, in particular, on requirements for transparency, documentation, and cybersecurity.

In terms of its impact upon financial services, the EU AI Act identifies two specific high-risk use cases, as indicated under Annex III Point 5 b) and c) of the Act, to which most obligations are imposed by the AI Act. These high-risk use cases are:

- **AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score**, with the exception of AI systems used for the purpose of detecting financial fraud.

- **AI systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance**.

The EU AI Act also prohibits certain activities where the use of AI systems allows to:

- Deploy subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques to materially distort the behaviour of a person or group of persons by appreciably impairing their ability to make an informed decision.

- Use an AI system that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation.

- Evaluate or classify natural persons based on their social behaviour or personal characteristics to create a "social score" leading to detrimental or unfavourable treatment.

- Create or expand facial recognition databases through untargeted scraping of facial images from the internet or closed-circuit TV footage.

The AI Act also affirms that extraterritorial scope applies when outputs of an AI system deployed outside of the EU are used in services by EU firms and consumers. Thus, Swiss or UK financial entities using such tools in relation to EU consumers fall under the AI Act's requirements.

Defining AI

The surveyed members highlighted a range of approaches to defining AI in their respective jurisdictions for the purposes of financial regulation. While most respondents indicated that a formal definition of "artificial intelligence" for financial market regulation already exists in their jurisdiction, some authorities

have defined AI elsewhere, but not specifically for the purposes of financial regulation, and others do not yet have a definition but are currently working on one. Key examples include:

- Eight jurisdictions who are EU member states adopted the definition provided under the EU AI Act, which is a supranational legislative instrument with direct applicability to EU member states (see Box 5 for more information).[44]
- Two other jurisdictions have defined the term under financial market regulation.[45]
- Three jurisdictions have adopted definitions provided under national instruments that are not specific to financial market regulation.[46]
- Five jurisdictions have not yet defined Artificial Intelligence specifically for the purposes of financial market regulation.[47]

Overall (and where defined) the definitions of AI used in the IOSCO member states show a high degree of commonality. This includes an emphasis on machine-based systems capable of autonomy and adaptiveness, the ability to process inputs and generate outputs (predictions, recommendations, decisions), the simulation or augmentation of human intelligence and applicability across a range of technologies, from deterministic (rule-based) to non-deterministic (learning-based) systems. Regulators' definitions often converge on international standards (e.g., the OECD definition or for EU members the EU AI Act).

Where authorities are taking definitions from international standards there is a high degree of commonality. For instance, both the EU AI Act and OECD frameworks aim to establish a common understanding of what constitutes AI while ensuring that governance focuses on systems with meaningful impact on people and society. A shared conceptual baseline supports coherence across international policy frameworks by grounding regulation in the core capacities and effects of AI systems rather than specific techniques.

Sandboxes and other initiatives

Several regulators have undertaken initiatives and other practices to approach AI uses by participants within the financial sector. These have taken the form of regulatory sandboxes, innovation hubs, or publications of principles that

[44] AMF (France), BaFin (Germany), HCMC (Greece), CONSOB (Italy), CBI (Ireland), CNMV (Spain), FI (Sweden), and AFM (The Netherlands).

[45] SEBI (India), and MAS (Singapore).

[46] SEC (USA), CSRC (China) and FSA (Japan).

[47] ASIC (Australia), CVM (Brazil), CMA (Saudi Arabia), FINMA (Switzerland), and FCA (United Kingdom).

specifically target such uses, with the aim of encouraging responsible adoption of AI technology. The survey found that 79% of respondents have implemented, or plan to implement, sandboxes or similar frameworks for AI evaluation.

Sandboxes are aimed at allowing regulators to assess the market impact and regulatory implications of AI while enabling firms to test new applications under supervised conditions. Sandboxes also aim to facilitate dialogue between regulators and firms, provide targeted supervisory guidance, provide early identification of risks, and give greater regulatory clarity. Participation rates and project numbers vary by jurisdiction, but the overall trend is toward broader adoption of sandboxes.

In particular, the UK FCA launched its AI Lab initiative in 2024[48], with the aim to support safe, responsible AI innovation in financial services, bringing together firms, academics and stakeholders to explore real-world AI use cases and risks, inform regulatory approaches, and help firms test and develop AI tools in appropriate testing environments. Its components include AI Live Testing, the Supercharged Sandbox, AI Spotlight, AI Sprints, and the AI Input Zone, which enables stakeholders to provide input into the FCAs understanding of AI use cases and risks. Notably, the FCA's Supercharged Sandbox, delivered in partnership with NVIDIA, provides an experimentation environment which lets financial firms test early-stage AI concepts, enterprise-grade tools, enrich datasets, and allows for direct regulatory engagement with participants, with the aim of helping firms build, refine and validate AI models in a secure, supervised setting prior to moving toward further deployment. AI Live Testing (AILT) is a component of the FCA's AI Lab that supports firms which are further along in development and are ready to begin deploying AI systems in UK financial markets. It enables firms to test AI systems in controlled live market environments, including where appropriate, interaction with real clients, while engaging with the FCA to better understand and mitigate potential risks to consumers and markets. AI Live Testing is designed as a collaborative and exploratory service rather than a supervisory or enforcement exercise and focuses on how safe and responsible AI can be implemented in practice as systems move into live use.

With regards to the publication of principles, Singapore's MAS published its Fairness, Ethics, Accountability and Transparency (FEAT) principles in 2018,[49] in order to guide financial firms in their use of AI and data analytics, and strengthen governance and trust to accompany the growth in the adoption of the technology, with a view to promote ethical, unbiased, accountable and transparent practices industry-wide. Building upon the FEAT principles, MAS, in partnership with the

[48] See: AI Lab | FCA.

[49] See: Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector.

industry through a consortium called MindForge,[50] has published the first of two AI risk management handbooks in 2025 that help financial institutions establish robust AI governance frameworks and enhance AI safety practices. This initiative complements MAS' supervisory AI Risk Management Guidelines by translating supervisory guidance on AI into operational practices that financial institutions can apply in a proportionate and risk-based manner.

Authorities have also undertaken initiatives to gain an understanding of the nature and scale of AI uses within their jurisdiction. At a European level for instance, a number of National Competent Authorities (NCAs) have participated with ESMA in a joint fact-finding exercise on the uses of AI within EU markets. Results are published in ESMA's first Trend Risk and Vulnerabilities Report (TRV) for 2026 and an ad-hoc TRV Risk Analysis, with certain NCAs also having published the results obtained within their jurisdiction. The AI Act will require the establishment of AI regulatory sandboxes throughout the EU in the near future.

Governance and risk management

As with many of the key areas of oversight, the Survey results highlighted that most responding IOSCO members supervise how financial service providers govern their use of AI across the AI system lifecycle, with about half of the respondents noting bespoke frameworks (largely the EU countries subject to the EU AI Act), while others have governance requirements that are part of general frameworks.[51] Many of those countries that rely on existing frameworks have issued guidance noting how these frameworks apply to AI-related governance.[52]

Of those respondents that do not currently have applicable frameworks, some are considering proposals to amend or add new regulations to address AI-related governance, for instance:

- China's CSRC has proposed a draft amendment to its Cybersecurity Law proposes adding framework provisions on AI safety and development.
- India's SEBI has proposed a framework of guiding principles to ensure ethical, transparent and accountable use of AI and ML by market participants.

50 See: Project MindForge.

51 ESMA statement on the use of AI in the provision of retail investment services, refers to governance requirements under existing EU regulations: ESMA provides guidance to firms using artificial intelligence in investment services.

52 OSC & QAMF (Canada): Staff Notice 11-348; SFC (Hong Kong): Circular on the use of generative AI language models; FSA (Japan): Principles for Model Risk Management; FINMA (Switzerland): FINMA Guidance 08/2024.

- Singapore's MAS has issued a consultation paper proposing a set of Guidelines on AI Risk Management to guide financial institutions on the responsible use of AI in the financial sector.

Risk management frameworks

The Survey results highlighted that for those jurisdictions with specific AI-related requirements, risk management frameworks had the following main elements of risks that relevant frameworks focused on addressing:

- **Inaccurate or fabricated results (hallucination risk):** Addressing the risk that GenAI systems may lead to inaccurate or fabricated results.
    - CSRC (China) required providers of GenAI services to take effective measures to improve the quality, authenticity, and accuracy of training data, aiming to reduce the risk of GenAI systems producing inaccurate or fabricated information. There are also prohibitions on using deep synthesis services to create or disseminate false news.[53]
    - The EU AI Act requires high-risk AI systems to implement robust data governance, risk management, and transparency measures to minimize inaccuracies and hallucinations. Providers must ensure data quality, conduct ongoing monitoring, and enable human oversight to detect and mitigate hallucinations, with additional obligations for general-purpose models to document risks and maintain technical transparency.[54]
    - SFC (Hong Kong) required firms to implement comprehensive frameworks for model validation, ongoing review, and monitoring of AI systems. For high-risk use cases, human-in-the-loop mechanisms and robust output testing are mandated to ensure outputs are accurate and not misleading, especially for client-facing generative AI applications.
- **Unfair bias:** Addressing the risk that financial service providers' use of GenAI systems may lead to unfairly biased outcomes.
    - Canadian Securities Administrator (CSA) (Canada) highlighted in a Staff Notice the importance of managing conflicts of interest in the context of using AI systems. Conflicts of interest that may be particularly salient where AI systems are used include the use of non-objective inputs that could result in biased recommendations or decisions detrimental to certain clients based on their demographic characteristics or in favor of proprietary products without due consideration of alternatives.

[53] https://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm.

[54] Regulation - EU - 2024/1689 - EN - EUR-Lex.

    - CSRC (China) requires firms to conduct fairness assessments of algorithms and prohibited them from setting discriminatory terms based on geographical or demographic differences. Effective measures must be taken to prevent discrimination based on ethnicity, belief, nationality, region, gender, age, occupation, health, and other factors during the processes of algorithm design, training data selection, model generation and optimization, and service provision.
    - Hong Kong Monetary Authority (Hong Kong) requires that AI-driven decisions must not discriminate or unintentionally exhibit bias, and that outcomes be fair, ethical, and consistent. It also required human-in-the-loop or manual intervention, and, specifically for early-stage client-facing GenAI, provide channels for clients to opt out or request human intervention and review.
    - The EU AI Act requires high-risk AI systems to implement data governance practices that detect, prevent, and mitigate unfair bias, ensuring datasets are representative and free of discriminatory effects.
- **Explainability risk:** Addressing the risks that may result from the use of AI systems that lack explainability.
    - Canadian Securities Administrator (CSA) (Canada) highlighted in a Staff Notice the importance of maintaining explainability and transparency in the context of using AI systems. When selecting or developing AI systems, the need for advanced capabilities of an AI system should be balanced with the need for explainability.
    - CSRC (China) required firms to disclose the algorithm logic, data sources, and decision-making basis, and to explain the interpretability mechanisms.
    - The EU AI Act's transparency concept is described as ensuring people are made aware when they are interacting with an AI system, while deployers (and affected persons) are informed about what the system can and cannot do and what rights or safeguards apply in that context. In addition, the EU AI Act introduces explicit expectations for appropriate human oversight measures for high-risk AI.
    - Hong Kong Monetary Authority (Hong Kong) required banks to ensure that AI applications are explainable "no black-box excuse" to all relevant parties, and to build adequate measures into the design phase to achieve a level of explainability that is appropriate and commensurate with the materiality of their AI application

Disclosure

Most IOSCO Member Survey respondents indicated that they currently regulate how issuers and/or supervised firms disclose their use of, or investment in, AI technologies, either through existing or bespoke frameworks.[55]

Some jurisdictions rely on the disclosure obligations in their existing regulatory frameworks and apply those obligations to the use of AI systems in capital markets. These frameworks typically encompass requirements for issuers and other market participants to provide accurate, complete, and timely information on material matters, including technological developments that may affect their operations, risk profile, or financial condition.[56]

Certain IOSCO members have developed bespoke regulatory regimes that include disclosure requirements in cases where AI systems are used in capital markets. For those jurisdictions, certain elements identified include:

- The risks, limitations, and consequences associated with the use of AI systems, including what data is used to make AI-driven decisions about the clients, how AI uses personal data and maintains confidentiality, the impact that AI may have on products or service, the consequences that AI-driven decisions may have on clients, and that the output generated may not be accurate.
- The use of AI in client interactions (e.g., chatbots, automated systems).
- The details about third-party service providers involved in AI systems, including their role and the nature of their technology.
- The technical details of AI algorithms, application scenarios, data sources, data collection methods, and data quality control.
- The use of AI systems for emotion recognition or biometric categorization, to protect individuals who may be subject to such technologies unknowingly.
- The use of AI systems if relevant to the investment strategy of the firm, and more broadly the role of AI systems in investment decisions

Notable examples of bespoke disclosure requirements or guidelines from the survey results include:

[55] CSRC (China), AMF (France), BaFin (Germany), HCMC (Greece), SEBI (India), CBI (Ireland), CONSOB (Italy), CNMV (Spain), FI (Sweden), AFM (The Netherlands), and, under consultation, MAS (Singapore).

[56] ASIC (Australia), OSC & QAMF (Canada), SFC (Hong Kong), and SEC (US).

- CSRC (China): Financial institutions must disclose in a legal, compliant, accurate, and timely way information about the AI algorithms they use to provide financial products and services.[57]

- European Union: Although the disclosure framework applicable to the provision of investment products and services in the EU is technology neutral, the EU AI Act contains additional disclosure requirements in specific cases, including:

    - Providers and deployers must ensure users are aware when they are interacting with an AI system regardless of the risk level. For example, providers of GenAI outputs must implement technical solutions to label AI-generated outputs in a machine-readable format, making it clear that the content was created by an AI system.

    - High-risk AI systems must be designed to ensure their outputs are interpretable by deployers, and must be accompanied by clear, complete, and accessible instructions for use.

    - Deployers must also verify that third-party AI systems they acquire comply with these standards.

    - Providers of general-purpose AI (GPAI) models must prepare and maintain extensive technical information and documentation relating to the training and functionality of such GPAI models. Providers of GPAI models will be expected to disclose different degrees of information to: (i) the AI Office and national competent authorities (upon request); (ii) other downstream providers of AI systems; and (iii) the wider public.

- India's SEBI: Bespoke regulations require an investment adviser to disclose to a client the extent of any use of AI systems in providing investment advice.[58] Disclosure must be made, by the entity, at the time of entering into a client agreement and make such additional disclosure

[57] See the Evaluation Specification of Artificial Intelligence Algorithm in Financial Application, promulgated by the People's Bank of China on March 26, 2021, and See the Administrative Provisions on Deep Synthesis in Internet-based Information Services, promulgated on November 25, 2022, by Cyberspace Administration of China, Ministry of Industry and Information Technology, and Ministry of Public Security.

[58] Regulation 15(14) and 18(9) SEBI (Investment Advisers) Regulations.

whenever required. Similar provisions exist for research analysts and research entities.[59]

- Singapore's MAS: MAS has published non-binding guidance focused on use of AI technologies that include the FEAT (Fairness, Ethics, Accountability, Transparency) principles. These principles encourage financial institutions to adopt transparent practices around AI system use.[60]

Recordkeeping

Where AI-specific recordkeeping requirements exist or are emerging, [61] they generally focus on AI systems considered "higher risk" by specific IOSCO members[62] and may include obligations to record events throughout the system's lifetime, document training data, algorithmic model, and risk management processes.

However, in some jurisdictions, current recordkeeping obligations are technology neutral as part of broader general regulatory record-keeping frameworks.[63] As an example, EU jurisdictions rely on different layers of recordkeeping requirements:

i) MiFID II (general framework): Firms are expected to keep records detailing how AI technologies are used in different areas related to the delivery of investment services. These records should cover all aspects of AI implementation, such as decision-making procedures, data sources utilized, algorithms applied, and any changes made over time.[64]

[59] Regulation 24(7) and 18(vii) of SEBI (Research Analysts) Regulations.

[60] Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector: FEAT Principles Final.pdf.

[61] Ten jurisdictions had record-keeping requirements that were part of bespoke frameworks, in addition to general frameworks: CSRC (China), SEBI (India) and European authorities, i.e. CBI (Ireland), CONSOB (Italy), CNMV (Spain), FI (Sweden), AFM (The Netherlands). AMF (France), BaFin (Germany), and HCMC (Greece).

[62] We note that at this stage, there is no consensus on what constitutes a high-risk AI system, and such concept may differ in each jurisdiction.

[63] ASIC (Australia), CVM (Brazil), OSC & QAMF (Canada), SFC (Hong Kong), FSA (Japan), and MAS (Singapore).

[64] ESMA Public Statement: On the use of Artificial Intelligence (AI) in the provision of retail investment services.

ii) DORA (general framework): Under DORA, AI systems are considered as Information and Communication Technology (ICT), especially when used for critical functions in financial services. DORA requires financial entities to report major ICT incidents within strict timelines and to notify clients of incidents affecting their financial interests.

iii) EU AI Act (bespoke framework): High-risk AI systems shall be able to automatically record events throughout their operational life. These logging features should ensure traceability by capturing events relevant to identifying risks, supporting post-market monitoring, and overseeing system operation.

Other jurisdictions rely on guidelines to clarify how the use of AI systems fits within their existing frameworks:

i) CSA (Canada) published guidance to clarify how the technology-neutral securities framework applies to use of AI systems by market participants.[65] Registered firms are required to maintain records to demonstrate the extent of the firm's compliance with applicable requirements under securities law, including KYC requirements and suitability determinations. AI systems used by registrants should provide an appropriate degree of explainability so that registered firms are able to meet applicable record keeping requirements.

ii) FSA (Japan) published several specific guidelines related to the use of AI by market participants. Firms are required to identify and inventory all AI models, with comprehensive documentation of methodologies, assumptions, and limitations. During the development process, firms should develop a comprehensive transparent model document detailing methodologies, assumptions, limitations, and weaknesses underlying the model. Firms are also required to assign a risk rating to each of the models.

iii) SFC (Hong Kong) published a specific guideline to set out expectations on licensed firms in respect of their use of generative AI. Licensed firms are required to document model testing, calibration, validation, and ongoing monitoring of AI systems - following a risk-based manner. For algorithmic trading, whether or not AI is used, firms must keep records of design, development, deployment, and operation of its electronic trading system.

[65] CSA Staff Notice and Consultation 11-348: Applicability of Canadian Securities Laws and the use of AI systems in capital markets.

Reporting

The Survey queried reporting requirements/regulation (supervision) on firm's AI incident reporting, firm's use of AI, and firm's investment in AI.

The majority of IOSCO member respondents highlighted that their respective jurisdictions have regulations that cover AI reporting, with the majority of those indicating that AI reporting regulation is part of the existing, technology-neutral reporting frameworks and were not specific to AI. A minority of responding members stated that there are specific AI reporting frameworks in their jurisdiction.

AI-related reporting is guided by broadly applicable technology-neutral requirements. For example, in the context of consumer protection, companies may have a duty of care to provide fair, clear and non-misleading information to clients. The explainability of the AI system is particularly relevant to meeting such requirements.

AI incident reporting

A key area of reporting is incidents involving AI. Responding IOSCO members to the Survey highlighted that AI incident reporting is well covered, either through specific AI incident reporting frameworks or through general ICT incident reporting.[66]

Based on the Survey responses received, outsourcing and third-party dependency information are covered by general reporting in jurisdictions that provided a response. Members reported that they require information on outsourcing and identification of critical third-party providers in relation to regulated activities mainly as part of technology-neutral existing regulatory frameworks. For example, firms subject to DORA in the EU are required to maintain and submit yearly, to their competent authority, a detailed "register of information", which is a comprehensive record of all their contractual arrangements with third-party ICT providers.

Specific AI-related information reporting

[66] AI incident reporting provides some information in relation to cybersecurity risks and fraud concerns in firms. These risks are identified by the 2025 IOSCO AI in Capital Markets report. For example, in the EU, digital operational resilience is required by DORA. An AI incident should be reported, under the DORA Regulation, if it qualifies as major ICT incident.

Most IOSCO members that responded to the Survey highlighted that they do not have an AI-specific legal framework, given AI use is a growing area, and therefore do not yet require specific AI related reporting on the use of AI or investment in AI.

For example, in the EU investment firm sector, there are general reporting requirements for firms using trading algorithms under the Markets in Financial Instruments Directive[67] (MiFID). Such entities reporting algorithms may utilize AI in their trading activities.

For IOSCO members that responded they have specific AI reporting, reporting generally covers:

- Functions within the firm where traditional AI and/or generative AI were deployed, whether the AI tool was developed in-house or by third party, as well as the use cases within the investment/risk management process in which generative AI was deployed (Hong Kong funds sector).

- Algorithm registration, amendment, and deregistration (China).

- Purpose and scope of the AI system; risk classification; incident details and mitigation measures; use of personal data and algorithmic decision-making (The Netherlands).

- Market infrastructure institutions submitting a reporting template on their AI use (India).

Applying existing regulatory frameworks to AI

Some respondents reported that there are existing guidelines in their jurisdictions on how to apply technology-neutral regulations to AI. For example, in December 2024, the Canadian Securities Administrators published their guidance on the "Applicability of Canadian Securities Laws and the Use of Artificial Intelligence Systems in Capital Markets"[68] to provide clarity and guidance on how securities legislation applies to the use of AI systems by market participants.

[67] Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EUText with EEA relevance.

[68] CSA Staff Notice and Consultation 11-348 – Applicability of Canadian Securities Laws and the Use of Artificial Intelligence Systems in Capital Markets.

Third-party and outsourcing risk

Based on the latest Survey responses, most jurisdictions apply existing requirements on outsourcing and governance of third parties to financial institutions using third-party products or services that rely on AI systems as opposed to a bespoke set of rules. Such existing requirements vary across jurisdictions but may broadly include risk management practices relating to responsibilities for senior management, risk assessments, confidentiality and security, ongoing monitoring and control of outsourcing arrangements. Some jurisdictions have issued guidance that supplements or clarifies existing outsourcing requirements in the specific context of AI. For example:

- In Hong Kong, the SFC has issued the Circular to licensed corporations - Use of generative AI language models.[69] The Circular mandates that licensed firms exercise due diligence and ongoing monitoring when using third-party AI service providers, assessing supply chain vulnerabilities and data leakage risks, with clear risk allocation in relation to cybersecurity, and establish appropriate contingency plans.

- In Canada, the CSA published Staff Notice and Consultation 11-348 Applicability of Canadian Securities Laws and the use of Artificial Intelligence Systems in Capital Markets.[70] Among other things, the Staff Notice states that registrants are responsible and accountable for all functions that they outsource to a service provider, must undertake due diligence before contracting for outsourced services, and must supervise any outsourced service provider on an ongoing basis. Ongoing supervision, in this context, will require registrants in some cases to review and verify samples of processes that use AI systems.

- In China, according to the Guidelines for Disclosure of Financial Application of Artificial Intelligence Algorithms, financial institutions are required to provide explanations for third-party software that is integrated with artificial intelligence algorithms and has the capability to collect and use personal information. When the tuning datasets are sourced from customized datasets provided by third-party data service providers or from databases in open-source communities, the source of the algorithm data must be disclosed. Additionally, during third-party external audits, the name and qualifications of the service provider conducting the compliance assessment of the artificial intelligence

69 Circular to licensed corporations - Use of generative AI language models | Securities & Futures Commission of Hong Kong.

70 CSA Staff Notice and Consultation 11-348 - Applicability of Canadian Securities Laws and the use of Artificial Intelligence Systems in Capital Markets | OSC.

algorithm application, the assessment time, the assessment conclusion, and the validity period of the assessment report should be explained.

Inspections and examinations

The results of the Survey highlighted that most responding IOSCO members include some level of oversight of AI use as part of their supervisory activity across sectors, while the specificity with which the oversight focuses on AI use or broader risk management, varies across IOSCO members. The majority include within their inspection programs, surveys, annual reporting, and thematic reviews questions to firms around how they are using AI and managing potential risks of AI use.[71] However some IOSCO members responded that AI is not yet a mandatory focus or is only covered under general principles.[72]

While some IOSCO members that include AI use as part of supervisory inspections cover all sub-sectors, some jurisdictions focus on more specific areas and use cases relating to AI identified through targeted inspections, including on areas such as portfolio valuation and management at asset managers (AMF), and robo-advice and portfolio management (AFM and AMF).[73]

Most IOSCO members do not mandate auditability mechanisms; however, some jurisdictions indicate auditability is required or operationally expected:[74]

- SFC (Hong Kong) requires firms to keep records of results of the GenAI model testing and calibration, validation and ongoing review and monitoring. Such information enables sampling for assessment.
- FINMA (Switzerland) does not mandate the use of auditability mechanisms in AI systems, but it is part of their expectation that auditability in AI systems is ensured.
- EU member states require conformity assessments for high-risk AI systems under the EU AI Act.[75]

[71] ASIC (Australia), OSC & QAMF (Canada), AMF (France), HCMC (Greece), SFC (Hong Kong), SEBI (India), CBI (Ireland), MAS (Singapore), CNMV (Spain), FI (Sweden), FINMA (Switzerland), AFM (The Netherlands), and FCA (United Kingdom).

[72] CVM (Brazil), CSRC (China), FSA (Japan), CMA (Saudi Arabia).

[73] AMF (France) and AFM (The Netherlands).

[74] CNMV (Spain), FI (Sweden), FINMA (Switzerland), and AFM (The Netherlands).

[75] See Box 5 for more information on the EU's AI Act.

Risk materiality assessments are increasingly expected. Seven IOSCO members indicated that risk materiality assessments are addressed through existing risk-based regulatory frameworks, including expectations around internal controls, governance, and model risk management. [76]

Most IOSCO members indicated that periodic reviews to detect degradation are required, and noted that they are addressed through existing risk-based regulatory frameworks, typically framed as risk-based monitoring, model governance, or outcomes-focused obligations.[77]

Likewise, most IOSCO members that conduct evaluations of AI use through their supervisory programs publish some form of information related to their findings, including:[78]

i. FSA (Japan) has published results from a survey on actual AI use by financial institutions and related entities, in an AI Discussion Paper.[79]
ii. MAS (Singapore) has published multiple AI-related public materials (presented as publications relevant to AI risk management and oversight).[80]
iii. FI (Sweden) has published a report on AI in the financial sector and risk management.[81]
iv. FINMA (Switzerland) has published various AI-related findings through its annual reports in 2021 and 2023, risk monitor 2023, one survey, and eventually as part of FINMA guidance.[82]
v. AFM (The Netherlands) has published AI-related findings through its robo-advice guidance and thematic reviews.[83]

[76] ASIC (Australia), CSRC (China), SFC (Hong Kong), FSA (Japan), FINMA (Switzerland), AFM (The Netherlands), and FCA (United Kingdom).

[77] SFC (Hong Kong), FSA (Japan), FINMA (Switzerland), and FCA (United Kingdom).

[78] See also FINRA's Annual Regulatory Oversight Report: GenAI: Continuing and Emerging Trends | FINRA.org.

[79] https://www.fsa.go.jp/news/r6/sonota/20250304/aidp.pdf.

[80] Information Paper on Artificial Intelligence Model Risk Management; Information Paper on Cyber Risks Associated with Generative Artificial Intelligence; Information paper on Cyber Risks Associated with Deepfakes; Project MindForge Phase 1 white paper.

[81] https://www.fi.se/en/published/reports/reports/2024/ai-increasingly-common-in-the-financial-sector-but-risk-management-is-lagging-behind/.

[82] https://www.finma.ch/en/news/2025/04/20250424-mm-umfrage-ki; https://www.finma.ch/en/news/2024/12/20241218-mm-finma-am-08-24.

[83] https://www.afm.nl/~/profmedia/files/onderwerpen/roboadvies-sav/visie-roboadvies.pdf.

vi. FCA (United Kingdom) has published several public reports related to AI use in financial services.[84]

Other IOSCO members keep findings of individual examinations confidential but may have general information available on their website.[85]

International cooperation and private sector engagement

The integration of AI into financial markets is occurring globally. As financial institutions expand their AI deployment and use cases become more complex, sharing supervisory practices and fostering dialogue between regulators and industry stakeholders as appropriate will help promote responsible innovation and consistent risk management in an interconnected financial environment.

The following section draws on the Survey results to highlight insights on international cooperation and private sector engagement.

International Cooperation

The Survey indicates that nearly all responding authorities actively collaborate with other organizations to monitor AI developments, utilizing international fora, workshops, and working groups. These mechanisms have facilitated information sharing, joint risk assessments, and the development of AI-related frameworks. While most respondents consider these cooperative efforts productive, a minority note ongoing challenges in cross-border coordination, particularly due to differences in legal frameworks and constraints on data sharing.

Private Sector Engagement

Engagement with the private sector is an important component of effective supervision. The Survey results indicate that nearly all authorities maintain regular dialogue with market participants, technology vendors, and academia through mechanisms such as roundtables, workshops, targeted surveys, and innovation fora. The majority of authorities report positive and constructive interactions with industry, financial consumer protection organizations, academia, and technology providers, viewing ongoing dialogue as essential for bridging

[84] Bank of England & FCA AI Survey; AI Sprint summary; firms' treatment of customers in vulnerable circumstances.

[85] For instance the SEC: *Division of Examinations*, https://www.sec.gov/about/divisions-offices/division-examinations and the EXAMS Risk Alerts at SEC.gov | Risk Alerts.

governance gaps and supporting the development of practical regulatory responses. Only a minority of responding IOSCO members reported explicit challenges in stakeholder engagement, with the main issues being low participation in voluntary consultations and, in some cases, firms' reluctance to share detailed information due to concerns about regulatory action or competitive disadvantages.

Annex II: Timeline of IOSCO's work on AI

Year	IOSCO initiative	Focus
2021	FR06/2021 The use of artificial intelligence and machine learning by market intermediaries and asset managers: "The 2021 AI Report"	• Identified key potential risks related to AI use by market intermediaries and asset managers and included guidance to assist IOSCO members in supervising AI use by such firms. • Highlighted the potentially transformative nature of AI technologies relating to their use in capital markets, as well as the key risks with respect to governance and oversight; algorithm development, testing, and ongoing monitoring; data quality and bias; transparency and explainability; outsourcing; and ethical concerns.
2022-2024	Fintech Task Force and AI Working Group established	• Established the Fintech Task Force (FTF) in March 2022 to lead IOSCO's work developing, overseeing, delivering, and implementing IOSCO's regulatory agenda with respect to Fintech. • Subsequently established the FTF AI Working Group (AIWG) in 2024 to focus on AI issues in a two phased approach.
2025	**Phase 1** CR/01/2025 Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges: "The 2025 AI Report"	• Outlined a shared understanding among IOSCO members of the issues, risks, and challenges presented by then-emerging AI technologies used in financial products and services, seen through the lens of investor protection, market integrity, and financial stability. • Identified how some IOSCO members were responding to those emerging developments, and steps being taken by market participants, in particular the emergence of GenAI and the potential impacts of this non-deterministic AI technology.
Current (H1 2026)	**Phase 2: Tranche 1** Supervisory Toolkit on AI (This report)	• Published a toolkit of potential supervisory approaches to evaluating the use of AI systems in capital markets, taking into account different national or regional frameworks and approaches.
From H2 2026	**Phase 2: Tranche 2** Review of emerging industry practices	• Plan to publish a review of emerging industry practices on the Disclosure, Reporting, and Governance of the Use of AI Systems in capital markets.

Annex III: List of IOSCO Reports that Discuss AI

- Update to the Report on the IOSCO Automated Advice Tools Survey (December 2016)
- IOSCO Research Report on Financial Technologies (FinTech) (February 2017)
- The use of artificial intelligence and machine learning by market intermediaries and asset managers (September 2021)
- Principles on Outsourcing (October 2021)
- The Use of Innovation Facilitators in Growth and Emerging Markets (July 2022)
- Report on Retail Distribution and Digitalisation (October 2022)
- Retail Market Conduct Task Force Final Report (March 2023)
- Policy Recommendations for Crypto and Digital Asset Markets (November 2023)
- Investor Education on Crypto Assets (October 2024)
- Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges (March 2025)
- Digital Engagement Practices (DEPs) (May 2025)
- Online Imitative Trading Practices: Copy Trading, Mirror Trading, Social Trading (May 2025)
- Thematic Review Assessing the Implementation of IOSCO Recommendations for Crypto and Digital Asset Markets (October 2025)
- IOSCO announces Call for Applications for its first TechSprint on Investor Education in the Age of Artificial Intelligence (AI) (March 2026)