The use of artificial intelligence and machine learning by market intermediaries and asset managers

Consultation Report

The Board
OF THE
INTERNATIONAL ORGANIZATION OF SECURITIES COMMISSIONS

CR02/2020	**JUNE 2020**

Copies of publications are available from:
The International Organization of Securities Commissions website www.iosco.org

Foreword

The Board of the International Organization of Securities Commissions (IOSCO) has published this Consultation Report to assist IOSCO members in providing appropriate regulatory frameworks in the supervision of market intermediaries and asset managers that utilise AI and ML.

How to Submit Comments

Comments may be submitted by one of the three following methods **on or before 26 October 2020**. To help us process and review your comments more efficiently, please use only one method.

Important: All comments will be made available publicly, unless anonymity is specifically requested. Comments will be converted to PDF format and posted on the IOSCO website. Personal identifying information will not be edited from submissions.

1. Email

- Send comments to consultation-02-2020@iosco.org.
- The subject line of your message must indicate '*The use of artificial intelligence and machine learning by market intermediaries and asset managers*
- If you attach a document, indicate the software used (e.g., WordPerfect, Microsoft WORD, ASCII text, etc) to create the attachment.
- Do not submit attachments as HTML, PDF, GIFG, TIFF, PIF, ZIP or EXE files.

2. Facsimile Transmission

Send by facsimile transmission using the following fax number: + 34 (91) 555 93 68.

3. Paper

Send 3 copies of your paper comment letter to:

Alp Eroglu
International Organization of Securities Commissions (IOSCO)
Calle Oquendo 12
28006 Madrid
Spain

Your comment letter should indicate prominently that it is a '*Public Comment on The use of artificial intelligence and machine learning by market intermediaries and asset managers'*

Contents

Chapter 1 - Executive Summary

Background

Artificial Intelligence (AI) and Machine Learning (ML), collectively called AI and ML, are increasingly being utilised in financial services, due to a combination of increased data availability and computing power. The use of AI and ML by market intermediaries and asset managers may be altering firms' business models. For example, firms may use AI and ML to support their advisory and support services, risk management, client identification and monitoring, selection of trading algorithms and portfolio management, which may also alter their risk profiles.

The use of this technology by market intermediaries and asset managers may create significant efficiencies and benefits for firms and investors, including increasing execution speed and reducing the cost of investment services. However, this use may also create or amplify certain risks, which could potentially have an impact on the efficiency of financial markets and could result in consumer harm. The use of, and controls surrounding AI and ML within financial markets is therefore a current focus for regulators across the globe.

IOSCO identified the use of AI and ML by market intermediaries and asset managers as a key priority. The IOSCO Board approved a mandate in April 2019 for Committee 3 on Regulation of Market Intermediaries (C3) and Committee 5 on Investment Management (C5) to examine best practices arising from the supervision of AI and ML.[1] The committees were asked to propose guidance that member jurisdictions may consider adopting to address the conduct risks associated with the development, testing and deployment of AI and ML.

Potential risks identified in the consultation report

IOSCO surveyed and held round table discussions with market intermediaries and conducted outreach to asset managers to identify how AI and ML are being used and the associated risks. The following areas of potential risks and harms were identified in relation to the development, testing and deployment of AI and ML:

- Governance and oversight;
- Algorithm development, testing and ongoing monitoring;
- Data quality and bias;
- Transparency and explainability;
- Outsourcing; and
- Ethical concerns.

Proposed IOSCO Guidance

This consultation report proposes guidance to assist IOSCO members in providing appropriate regulatory frameworks to supervise market intermediaries and asset managers that utilise AI and ML.

[1] Board Priorities - IOSCO work program for 2019, March 25, 2019, available at: https://www.iosco.org/library/pubdocs/pdf/IOSCOPD625.pdf

The proposed guidance consists of six measures that reflect expected standards of conduct by market intermediaries and asset managers using AI and ML. Although the guidance is not binding, IOSCO members are encouraged to consider these proposals carefully in the context of their legal and regulatory frameworks. IOSCO members and firms should also consider the proportionality of any response when considering these proposals.

The use of AI and ML will likely increase as the technology advances, and it is plausible that the regulatory framework will need to evolve in tandem to address the associated emerging risks. Therefore, this report, including the definitions and guidance, may need to be reviewed in the future to remain up to date.

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 requiring firms to have 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;

(b) operate in a way that complies with regulatory obligations.

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 sanctions 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.

Chapter 2 - Background and Scope

Previous IOSCO work in this area

IOSCO has undertaken several workstreams on the use of AI and ML in financial markets, including:

- **Committee on Emerging Risks (CER):** The CER undertook a mandate on the use of novel technologies deployed by regulators to increase the efficiency and effectiveness of supervisory and oversight programs and published a report in February 2017.[2] CER examined the regulatory use of tools such as big data analytics and data visualisation technologies; AI and ML, and deep learning technologies; and distributed ledger technologies.

- **Committee on Regulation of Secondary Markets (C2):** C2 published a report in April 2013 on Technological Challenges to Effective Market Surveillance Issues and Regulatory Tools.[3] The report made recommendations to help market authorities address the technological difficulties facing effective market surveillance.

- **IOSCO Fintech Network:** The IOSCO Fintech Network was established in May 2018 to facilitate the sharing of knowledge and experiences among IOSCO members. The IOSCO Fintech Network considered the ethical implications of the use of AI and ML technologies.

IOSCO Mandate

Building on the previous IOSCO work, the proposed guidance seeks to address the potential risks and harms that may be caused by the use of AI and ML by market intermediaries and asset managers and looks to help ensure that market intermediaries and asset managers have:

- appropriate governance, controls and oversight frameworks over the development, testing, use and performance monitoring of AI and ML;
- staff with adequate knowledge, skills and experience to implement, oversee, and challenge the outcomes of the AI and ML;
- robust, consistent and clearly defined development and testing processes to enable firms to identify potential issues prior to full deployment of AI and ML; and
- appropriate transparency and disclosures to their investors, regulators and other relevant stakeholders.

---

2 *IOSCO Research Report on Financial Technologies (Fintech)*, February 2017, available at: https://www.iosco.org/library/pubdocs/pdf/IOSCOPD554.pdf

3 *Technological Challenges to Effective Market Surveillance Issues and Regulatory Tools*, August 2012, available at: https://www.iosco.org/library/pubdocs/pdf/IOSCOPD389.pdf

Defining the terms Al and ML for this consultation report

AI can be understood as a combination of mass data, sufficient computing resources and machine learning. ML is a sub-set of AI which can be defined as a method of designing a sequence of actions to solve a problem, which optimise automatically through experience – with or without human intervention.

Artificial Intelligence

The term "Artificial Intelligence", first coined by data scientist John McCarthy[4] in 1956, can be understood as a combination of mass data, sufficient computing resources and ML, which can accomplish simple, repetitive tasks, or can be more sophisticated and, to some degree, self-learn and perform autonomously, based on a system that mimics human cognitive skills or human capabilities. However, the prospect of a computer having such a level of intelligence, also called "strong artificial intelligence" is not expected in the foreseeable future. AI in the financial services industry is still in its relative infancy and is poised to become more common, and with that will come legal, ethical, economic and regulatory challenges.

Machine Learning

The term "Machine Learning" is a specific subset and application of AI, which focuses on the development of computer programs that analyse and look for patterns in large quantities of data, with the aim of building knowledge to make better future decisions. ML algorithms differ from traditional algorithms in their ability to harness inductive reasoning, i.e., ML algorithms learn and develop using past data trends to predict future outcomes. Inductive reasoning is often used to predict future outcomes, with the accuracy of the hypothesis improving as more observations are made and the quality of the data improves. This however cannot be guaranteed due to potential data shortcomings.

A successful ML algorithm will therefore learn and evolve over time and will possibly make recommendations that were not explicitly envisaged when it was created. [5]

There are various categories of ML algorithms. These categories are based on the level of human intervention required in feeding back, and/or interpreting data from the algorithm. "Deep Learning", a form of computationally intensive ML that learns association and statistical patterns often from a very large dataset, can also include any of these categories: [6]

---

4 *What is AI?* available at: http://jmc.stanford.edu/artificial-intelligence/index.html

5 A famous example of this is the "move 37" in the game of Go: when Google's AlphaGo5 algorithm was pitted against professional Go player Lee Sedol in March 2016, making a move on the 37th turn that was previously unimaginable. This algorithm used "deep learning", a form of ML technique that efficiently learns associations and statistical patterns from a very large dataset.

6 Deep learning is a method that analyses data in multiple layers, starting with learning about simple concepts then learning more complex concepts. Deep learning can be used for all three categories of machine learning algorithms.

- Supervised learning: the algorithm is fed an initial set of data that has been labelled. Based on this training set, the algorithm will learn classification rules and predict the labels for the remaining observations in the data set.

- Reinforcement learning: the algorithm is fed an initial set of data that has not been labelled and is asked to identify clusters of observations underpinned by similar characteristics. As it chooses an action for the data points, it receives feedback that helps it learn.[7]

- Unsupervised learning: the algorithm detects patterns in the data by identifying clusters of observations underpinned by similar characteristics – it uncovers the structure of the data on its own.

[7] R Sutton, A Barto, Reinforcement Learning: an introduction, MIT Press, 1998.

Chapter 3 – How firms are using AI and ML techniques

AI and ML use by market intermediaries and asset managers

Market intermediaries and asset managers' use of AI and ML is growing, as their understanding of the technology and its utility evolves. The rise in the use of electronic trading platforms and the increase of available data have led firms to consider the use of AI and ML in several areas, including for their trading and advisory activities, as well as for risk management and compliance.

The IOSCO firm engagement revealed that within financial markets, AI and ML are being adopted to augment existing processes and activities, with a view to reducing cost and increase efficiency. AI and ML are freeing up resources to focus on more cognitive aspects, such as strategy, portfolio selection and generating investment ideas. Market intermediaries are deploying this technology in:

- Advisory and support services;
- Risk management;
- Client identification and monitoring;
- Selection of trading algorithm; and
- Asset management/ Portfolio management.

The use of AI and ML by asset managers appears to be in its nascent stages and is mainly used to support human decision-making. AI and ML techniques are being used to:

- Optimise portfolio management;
- Complement human investment decision-making processes by suggesting investment recommendations; and
- Improve internal research capabilities, as well as for back office functions.

Some asset managers are also beginning to use AI and ML for order execution, broker selection and order routing purposes (including through methods such as algo-wheels).[8]

Advisory and support services

Most robo-advisors or automated investment advisors use simple, rule-based (i.e., deductive) algorithms, although some are beginning to utilise predictive ML algorithms. Where ML is used to provide advisory services, most firms have manual intervention processes. The automated advice system is therefore usually limited to generating potential advice or asset allocation for the investment adviser to review. The investment adviser can then use this AI-generated advice as appropriate and where suitable, to make a recommendation to the client.

[8] We understand algo-wheels may perform different functions in different parts of the world. In this context, we understand algo-wheels to mean a software/model that aggregates data to select the strategy and broker through which to route orders before generating a report that sets out the reason behind how and where the trade was made.

Risk management

Risk management involves using data to price and manage exposure, including credit, market, operational and liquidity risk. Market intermediaries are harnessing ML based risk management systems for credit risk monitoring which could help provide an early-warning indicator of potential customer defaults and can help create a dynamic measurement of a customer's risk profile to better understand when to write off a debt.

ML is improving the efficiency of back office processing and reporting functions within market intermediaries. It is also increasingly used to visualise market risk by analysing volatility trends, gauge liquidity risk by analysing multi-dimensional risk and exposure data. Monitoring staff e-mails is increasingly being performed using ML algorithms by leveraging advanced pattern recognition.

Some market makers, who provide liquidity to market participants by facilitating transactions are adopting ML models and reinforcement learning to minimise their inventory risk and maximise the utility of their balance sheet.

Similarly, some asset managers are seeking to harness the advantages of these techniques in risk management. Some hedge funds and asset managers are automating risk management and compliance processes by tracking the behaviour of individual portfolio managers, automating execution quality reports and assessing market liquidity risk.

Client identification and monitoring

ML has allowed market intermediaries to automate their client onboarding, fraud detection, money laundering and cyber-attack monitoring. Market intermediaries must undertake Know Your Customer (KYC) checks before onboarding clients and selling them products and services. KYC entails collection and verification of comprehensive personal information from potential clients, involving processing unstructured metadata.

Inductive reasoning algorithms help accurately identify fake photo IDs, while recognising different photos of the same person. ML can also be used for screening and monitoring clients and transactions against sanctions or other lists, to detect possible money laundering, terrorist financing and other financial crimes.

Selection of trading algorithms

Many market intermediaries currently offer a software solution to their clients that selects an appropriate trading strategy and/or a broker depending on the market situation and trading objectives for best execution purposes, often named "algo wheel". Algo wheels seek to classify historical trading and performance, predict the performance of strategies and broker algorithms, and recommends when to use which algorithms. Using algo wheels to optimally execute simpler orders allows the trader to focus on more complex trade flows.

Predictive data analytics are enabling the identification of potential market conditions conducive to a "flash crash" type event.

Asset management/Portfolio management

Supervised learning, where a function is inferred from labelled training data, has been used for small-scale pattern recognition and simple prediction models to aid trading decisions within asset managers and market intermediaries for several years.

Margin pressure and competition however is driving innovation amongst asset managers. To compete, some active managers that traditionally emphasised their fundamental research capabilities are beginning to expand already existing quantitative approaches by leveraging diversified data sources – such as social media, geospatial data, and other metadata to enhance internal research.

In certain cases, these methods are being used for asset-allocation and pricing such as identifying relationships in metadata which could be used to generate trade ideas or "alpha signals" and forecast asset prices based on historical prices as well as current trends. Moreover, asset managers may apply these techniques to price new investment products competitively. They do so by using data on existing investment products with similar structures and/or constituent assets. Other applications include investment compliance checks, transfer agency activities and client servicing.

Chapter 4 – Identified Potential Risks and Harms Posed by the Use of AI and ML

IOSCO's industry engagement revealed that the evolution and increasing adoption of AI and ML may raise a number of conduct (either intentionally or unintentionally) concerns for market intermediaries and asset managers, including:

- Governance and oversight;
- Algorithm development, testing and ongoing monitoring;
- Data quality and bias;
- Transparency and explainability;
- Outsourcing; and
- Ethical concerns.

Governance and oversight

Firms implementing AI and ML mostly rely on existing governance and oversight arrangements to sign off and oversee the development and use of the technology. In most instances, the existing review and senior leadership-level approval processes were followed to determine how risks were managed, and how compliance with existing regulatory requirements was met. AI and ML algorithms were generally not regarded as fundamentally different from more traditional algorithms and few firms identified a need to introduce new or modify existing procedural controls to manage specific AI and ML risks.

Some firms indicated that the decision to involve senior leadership in governance and oversight remains a departmental or business line consideration, often in association with the risk and IT or data science groups. There were also varying views on whether technical expertise is necessary from senior management in control functions such as risk management. Despite this, most firms expressed the view that the ultimate responsibility and accountability for the use of AI and ML would lie with the senior leadership of the firm.

Some firms noted that the level of involvement of risk and compliance tends to focus primarily on development and testing of AI and ML rather than through the lifecycle of the model (i.e., implementation and ongoing monitoring). Generally, once implemented, some firms rely on the business line to effectively oversee and monitor the use of the AI and ML. Respondents also noted that risk, compliance and audit functions should be involved throughout all stages of the development of AI and ML.

Many firms did not employ specific compliance personnel with the appropriate programming background to appropriately challenge and oversee the development of ML algorithms. With much of the technology still at an experimental stage, the techniques and toolkits at the disposal of compliance and oversight (risk and internal audit) currently seem limited. In some cases, this is compounded by poor record keeping, resulting in limited compliance visibility as to which specific business functions are reliant on AI and ML at any given point in time.

Algorithm development, testing and ongoing monitoring

It is important that firms have robust and well understood development and testing frameworks in place, regardless of whether they are using AI and ML or traditional algorithms.

Overall, IOSCO's engagement showed that in most cases there is not an established framework for specifically developing AI and ML. Instead, many firms use the same development and testing frameworks that they use for traditional algorithms and standard system development management processes.

Firms that use algorithms should consider maintaining an appropriate development and testing framework, which is consistently applied across all relevant aspects of the business. This is particularly important where firms are using AI and ML within their algorithmic trading strategies.

Algorithms rely on quality data and most firms recognise the need for quality data inputs. Excessive immaterial, or "noisy" data is unwanted data that does not contribute to a relationship and may cause ML algorithms to miss the signal in the data and behave unexpectedly.

Robust development and testing controls are necessary to distinguish signals and statistically significant data from the noise. Unlike traditional algorithms, as more data is processed by ML algorithms, they may behave unpredictably as they are exposed to new data patterns. ML algorithms should therefore also be continuously monitored throughout their deployment to help ensure they do not behave inexplicably owing to a subtle shift in the operating conditions or excessive noise. While some firms noted they review and revise existing models as necessary, some firms focus less on managing models in the post-production phase so that they perform as they should over time. Unlike traditional algorithms, ML algorithms continually learn and develop over time. It is important that they are monitored to ensure that they continue to perform as originally intended.

Data quality and bias

The performance of AI and ML is inherently dependent on the quality of the dataset particularly when building the model. Learned bias in the dataset can impact the decisions made by such algorithms and may result in discriminative decisions and provide undesirable outcomes to market participants. For example, asking questions phrased in a certain way or in a certain sequence may lead to a response that introduces implicit or explicit bias from the respondents. Such a dataset, where a bias may have been introduced by either the questioner or by the respondents, will influence the conclusions reached by the algorithm. Any output based on such a bias will likely degrade the performance of the algorithm more quickly over time and could result in consumer detriment. [9]

[9] The French regulator Authorite des Marches Financiers (AMF) acknowledges that algorithms can reinforce biases and lead to financial exclusion of certain groups. For example, models based on behaviour history perform less well for younger customers with a limited history, in which case other explanatory variables need to be found. available at: https://acpr.banque-france.fr/sites/default/files/medias/documents/2018_12_20_intelligence_artificielle_en.pdf

Bias may also be inadvertently introduced during data cleansing – a pre-processing step often necessary to improve data quality. Cleaning data before applying ML can increase the signal-to-noise ratio allowing for more meaningful interpretations to be derived. However, cleansing data involves subjective decisions, which may inadvertently introduce other biases.

Transparency and explainability

The effective use and adoption of AI and ML require algorithms that are not only accurate but are also understandable by firms (including front line, compliance and risk personnel), market counterparties, clients and regulators. While increased transparency in firms' use of AI and ML could improve public understanding and confidence in the use of the technology, excessive transparency could create confusion or opportunities for individuals to exploit or manipulate the models. The level of transparency will also differ depending on the audience; for example, a regulator may require more detailed information than a client. These considerations need to be balanced in determining the appropriate level of transparency in the use of AI and ML.

It is important that firms appropriately disclose information about their service offerings, to help clients understand the nature and risks of products and service offerings, so that they can make informed decisions. Applying unexplainable ML algorithms to refine a trading strategy could expose the firm to unacceptable levels of legal and regulatory risk. Firms offering automated investment services, including "robo" advice and automated portfolio construction should appropriately disclose the nature of the use of ML and the automation in their service offering.

Some ML models operate as a "black box" with limited clarity on the reasoning behind the output. For example, in the case of deep unsupervised learning algorithms, the decisions made by the AI and ML models can be non-interpretable or unexplainable.[10] The interpretability or explainability of ML models on an ex-ante basis may therefore be challenging.

Firms had differing views on what is reasonably expected to be disclosed to investors. Some firms suggested that more transparency should be provided with respect to the use of AI and ML in investment decisions as opposed to administrative and back office functions. While most firms supported a level of generic transparency, others were against disclosures on the specific implemented approaches. Some firms also suggested there may be differences in disclosures provided to institutional versus retail clients.

Outsourcing

IOSCO's engagement observed that firms used external providers of AI and ML to varying extents. Some larger firms indicated they develop and implement AI and ML in-house or in partial collaboration with external providers, however smaller firms tend to revert to solutions offered by external providers. Many firms also use externally offered data pools and some firms noted they rely on external providers for cloud computing. Use of external providers for AI and ML solutions, including data pools and cloud computing may raise concerns about data privacy, cybersecurity and operational risk in addition to concerns about concentrations of certain providers (e.g., data providers, cloud providers, technology providers), particularly

[10] Basic ML models are explainable, where the more advanced ones (that use deep learning and have neural networks) may encounter the "black box" problem.

where firms may not have the appropriate level of expertise to conduct effective due diligence assessments ahead of utilising their services.

Ethical concerns

Since the financial crisis, significant attention has been paid to the role of ethics and trust in financial services by market participants. The Chartered Financial Analyst (CFA) Institute has, for example, defined ethics as "*a set of moral principles or rules of conduct that provide guidance for our behaviour when it affects others*" and has suggested that fundamental ethical principles include honesty, fairness, diligence, care and respect for others.[11] Ethical conduct follows those principles and balances self-interest with both the direct and indirect consequences of that behaviour for other people.

In the context of AI and ML, ethical concerns may arise when models develop certain social biases and potentially recommend undesirable outcomes, for example where data cleaning, data transformation and anonymisation of data were not adequately considered. There are questions about how these types of ethical considerations can continue to be met by firms and their employees as algorithmic models play an increasingly important role in the functioning of markets.

The IOSCO Fintech Network, sought to leverage existing academic literature and industry engagement to understand how the use of AI and ML may affect ethical behaviour in securities markets. In its engagement with industry participants and member organisations, the Fintech Network has identified robo-advice as one AI and ML application with significant potential ethical implications.

While most robo-advisors use simple rule-based algorithms, some are moving towards the use of predictive machine learning algorithms. Market participants should be careful when developing such AI and ML using large pools of alternative (i.e., non-traditional) datasets, such as satellite data or twitter feeds, to seek to ensure that the developed models would not discriminate against a certain segment of the population and that the AI and ML driven decisions are fair and unbiased.

In this context, IOSCO's Fintech Network identified five primary themes that could underpin the ethical use of AI and ML techniques:

- Beneficence – "do good": ensuring the model is being used and/or acting in good faith, in the best interest of investors and with market integrity.
- Non-malfeasance – "do no harm": having the ability to understand and interpret AI and ML based decisions to identify where misconduct may be taking place.
- Human autonomy, including auditability: ensuring humans have power over what the model can and cannot decide.
- Justice: accountability and transparency: ensuring there is accountability for the actions of the model at a senior level; and that accountability comes with appropriate understanding of the models. This theme is also about firms understanding the level of transparency they need to demonstrate both internally and with clients.

[11] https://www.cfainstitute.org/en/ethics/codes/std-of-practice-guidance/ethics-and-investement-industry

- Explainability: ensuring the outcomes arising out of the models used can be explained.

Many of these themes are explored, directly and indirectly, in Chapter 5 – Firm responses to the potential risks arising from the use of AI and ML and in IOSCO's proposed guidance.

Chapter 5 –Firms' responses to the potential risks arising from the use of AI and ML

Obligations under existing regulations

The AI and ML techniques described above are currently being deployed by some firms and used in the context of existing regulatory frameworks. This includes high level generic rules in relation to systems and controls, senior management accountability, risk management, outsourcing and governance. However, the rise in the use of these techniques has led to increased interest from regulators.

Many jurisdictions have overarching requirements for firms' overall systems and controls, however few jurisdictions have regulatory requirements that specifically apply to AI and ML based algorithms. These overarching requirements include rigorous testing of the algorithm in a safe environment away from the market before its controlled deployment, as well as continuous performance monitoring throughout the lifecycle of the algorithm. In this context, firms should consider whether solely relying on existing processes remains appropriate when meeting regulatory requirements or whether those processes should be further revised.

Moreover, not all ML techniques will be compatible with existing legal or regulatory requirements. Regulatory regimes that require algorithms to be fully understood and explainable throughout their lifecycle invariably limits the use of algorithms that evolve through the course of deployment in response to environmental changes.

Firms not using AI and ML

Some firms stated that they are opting not to use or are limiting their use of AI and ML algorithms until they are able to comply with jurisdictions' laws and regulations applicable to the use of the technology. Numerous survey respondents commented that they did not have the human resources or the right expertise at all levels to always fully understand AI and ML algorithms. Some firms therefore cannot implement the technology in a compliant manner. Operational and infrastructural concerns such as compatibility between AI and ML and their legacy technology, and operational processes are also precluding firms from meaningfully adopting AI and ML. Costs may be an additional consideration for small and medium-sized firms.

Potential mitigations to the risks posed by the use of AI and ML?

Through IOSCO's engagement with firms, several considerations were put forward as a way to mitigate the potential risks posed by the use of AI and ML, including ethical ones as identified by IOSCO's Fintech Network. These firms generally considered the following elements as pertinent to risk mitigation:

- Culture;
- Accountability;
- Knowledge/expertise/skills; and
- Operational resilience.

Culture

As the culture of each firm is different, a "one size fits all" approach would unlikely be effective. However, when considering a framework conducive to good market conduct for the use of AI and ML, the key elements of codes of ethics (e.g, honesty, fairness, diligence, care and respect for others) can provide the foundations to build upon. Senior leaders of a firm can manage the drivers and behaviours within the firm to create and maintain a culture, or "tone from the top", which can help reduce any potential harm to investors caused by inappropriate behaviour.

Firms also reported that the ethical implications from the use of AI and ML techniques could be viewed within this existing framework. Senior leaders can promote the organisation's culture and can help minimise the risk from their use of ML. Firms that are considering ethical frameworks for the use of AI and ML systems in the broader context of the organisation's culture could reduce the harm to investors by designing AI and ML systems and a control environment that reflect ethical values, organisational values and cultural behaviour.

This can be demonstrated by developing a robust governance framework and the availability of the right skills, knowledge, expertise and experience that is supported by a culture that permits adequate challenge and accountability across the firm. Responsibilities do not solely lie with senior leadership and second or third line functions, but also with the front office functions which act as the first line of defence with regards to the development and operations of AI and ML. Firms should also have consistent standards for developing and testing AI and ML algorithms. In this context, considerations about the appropriate level of disclosure to clients, regulators and all relevant stakeholders of where AI and ML techniques are used can also assist the firm in promoting a transparent culture.

Accountability

Many jurisdictions indicated that they have recently considered whether and how individual accountability can promote fairness and trust within financial services. Accountability aims to reduce harm to investors and strengthen market integrity by making individuals personally accountable for their conduct and competence. For example, under the UK's Senior Managers & Certification Regime (SM&CR), senior managers are ultimately accountable for the activities of the firm, and there is a certification requirement for staff responsible for algorithmic trading.

It is important that senior managers have adequate oversight and control over the development, testing, implementation and monitoring process for AI and ML techniques, as well as over any third-party providers to which their firm may be outsourcing.

Developments in AI and ML technologies may lead regulators and firms to further consider the responsibilities of other individuals, such as data analysts, data scientists and data engineers who may not typically have performed roles subject to regulatory scrutiny.

Knowledge/Skills/ Expertise

Ensuring the right skills, knowledge and expertise to discharge the responsibilities of an employee's role is essential. This includes standards of ethical behaviour.

Firms may want to consider whether to reassess their approach to evaluating individuals' skillsets and also ethical attributes, such as honesty, integrity and reputation, in order to meet the challenges posed by AI and ML. One firm specifically mentioned that it has educated all employees and senior management on the development and application of AI and ML techniques and how the use of this technology is consistent with the firm's culture. This has reportedly helped bridge the gap between technical and non-technical staff across the firm. Where firms outsource certain components of their technology, there may also be concerns about the level of understanding, at firm level, of the activities that take place at third-party provider level which can in turn impact the quality of due diligence reviews.

Operational resilience

Operational disruption can impact market integrity, threaten the viability of individual firms, or cause harm to consumers and other market participants. Should the use of AI and ML become widespread, operational resilience could also impact financial stability as new and unexpected forms of interconnectedness between financial markets and institutions emerge. As AI and ML are reliant on large data sets and computational resources, more firms are leveraging cloud computing offered by third-party service providers, who offer potentially cheaper, more secure and "ever-green" technology solutions. Many firms flagged operational resilience as one of their key concerns when using AI and ML techniques and mentioned their due diligence and oversight of third-party service providers as a key mitigating control.

Chapter 6 – Proposed Guidance

This chapter contains proposed guidance to IOSCO members. Each measure is designed to address one or more of the key potential risks and harms identified in Chapter 4. They are informed by good practices currently being carried out by some firms or expected by several regulators across the globe.

Although the proposed guidance is non-binding, IOSCO encourages its members to consider the guidance within the context of their legal and regulatory framework when regulating firms utilising AI and ML. The proposed guidance reflects an expectation of high standards of conduct across market intermediaries and asset management firms, and includes:

- Governance and responsibilities;
- The development, testing, and ongoing monitoring of AI and ML techniques;
- Knowledge and skills required by firms' staff;
- Operational resilience;
- Systems and controls; and
- Transparency and disclosure.

Proportionality

Proportionality should underpin the consideration of the proposed guidance. Regulators and firms should consider the risks posed by AI and ML while not undermining key safeguards that the guidance seeks to put in place. In judging proportionality consideration should be given to the activity that is being undertaken, the complexity of the activity, risk profiles, and the potential impact that the technology has on client outcomes and market integrity. Regulators should use their own judgement when applying this overarching principle.

When considering proportionality, it is not appropriate to consider the size of the firm as the only factor. In practice, the activities of a relatively small firm could have a material impact on market participants and the orderly functioning of markets. Regulators should consider what impact the technology could have on client outcomes, as well as markets. Client facing tools that rely on AI and ML, for example, might have a greater impact on client outcomes as opposed to where the technology is supporting back or middle office functions.

Given the significant potential risks and harms the measures intend to address, IOSCO members are encouraged to consider the extent to which these measures should be implemented in the context of regulated entities, business models, and their own legal and regulatory framework.

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 requiring firms to have 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.

This measure looks to embed accountability in all aspects of a firm's use of AI and ML and helps ensure the technology is appropriately understood, tested, deployed and monitored. The most senior people (senior management) in firms who perform key roles (such as "senior management functions" in the UK) are typically accountable for the overall performance of the firm. This accountability extends to the actions and outcomes of AI and ML models, including externally sourced models. This responsibility requires clear lines of accountability, including procedures to approve the development, deployment and subsequent updates of trading algorithms and to solve problems identified when monitoring trading algorithms. Clear lines of accountability include firms having a documented internal governance framework (which include the legal, compliance and risk management). Regulators should consider requiring firms to:

- Understand how AI and ML are being utilised, including the intended outcomes;
- Implement appropriate controls and governance frameworks to oversee and challenge the outcomes derived from the AI and ML models, including externally sourced models;
- Formulate a clear methodology document their methodology and have an audit trail of the use of AI and ML across the life cycle of the use of AI and across their business; and
- Assess whether the technology is applied consistently with the firm's risk appetite and client's risk tolerance, and in an ethical manner.

Senior management may already have the appropriate technical knowledge required to effectively oversee the firm's use of AI and ML techniques. However, where they do not have such knowledge, they could designate appropriate senior personnel within the firm to support them in discharging this oversight role while retaining ultimate accountability.

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;

(b) operate in a way that complies with regulatory obligations.

Any testing should reflect the underlying complexity and systematic risks posed by the use of AI and ML. Any material changes to the firms' systems should trigger further testing.

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.

Once deployed, the performance and output of the AI and ML algorithms should be monitored on a real-time basis. Firms should ensure adequate "kill switch" functionality is built into their

control framework, and that appropriate functions and people have access to it. “kill switch” functionality should also be appropriately tested.

The use of AI and ML should be properly assessed and tested in light of their risks, including market abuse, data privacy, risk management, and cybersecurity to help ensure they work as intended. Risk and compliance functions should be involved in the development and testing of AI and ML, as well as continuously monitoring the outputs post-deployment.

The behaviour of AI and ML may change in an unforeseen manner as more data is processed over time. Firms should therefore think beyond the existing testing and oversight arrangements that may be used for traditional algorithms and ensure the AI and ML techniques can be monitored continuously as the algorithms adjust and transform.

Therefore, it would not be enough for the AI and ML algorithm to be tested thoroughly before deployment; they need to be continuously monitored throughout their deployment to ensure that an algorithm does not behave in inexplicable ways owing to a subtle shift in the operating conditions or due to excessive noise. The “kill switch” function should also be complemented with appropriate back-up solutions and tested prior to live deployment to ensure that this function can indeed be depended on and initiated should it be required in future.

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.

The lack of adequate internal skills, knowledge, expertise and experience to maintain and oversee AI and ML may lead to difficulties in updating the model or over-reliance on external parties. It is therefore essential that internally, firms have the appropriate skills, knowledge, experience and expertise including front line, compliance, risk management and senior management staff, to understand and continuously supervise AI and ML within existing legal and regulatory parameters.

Firms should consider establishing multi-disciplinary teams involving the business line users of these solutions, data scientists, IT and database administration staff, as well as risk management and compliance teams.

Finally, firms should consider the importance of model continuity and mitigate the risk that the departure of key quantitative researchers or data scientists disrupt the proper functioning of the AI and ML solutions. Firms should therefore have processes and documentation requirements that provide for model continuity in the event of staff departures.

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 sanctions for poor performance.

Operational resilience and heavy reliance on third-party service providers is one of firms' key concerns when using AI and ML techniques. To ensure adequate accountability, firms should perform initial and ongoing due diligence and 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 key performance indicators and should also clearly determine suitable recourse sanctions for poor performance.

In addition, in recognition of the new challenges and risks involved in operating internationally or outsourcing significant level of activities to third parties, IOSCO is currently updating its existing principles on outsourcing to consider the risks associated with the use of third-party providers. IOSCOs outsourcing principles should be considered when outsourcing the development, testing and oversight of AI and ML.

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.

Firms should disclose meaningful information to customers and clients around their use of AI and ML algorithms that impact their outcomes. The objective is to disclose sufficient information to enable clients to understand the nature of, and key characteristics of the products and services that they are receiving, and how they are impacted by the use of the technology. Firms should consider the level of detail necessary to satisfy the disclosure objective. This transparency should aim to provide, on a non-discriminatory basis, all relevant parties with the necessary information to evaluate, to the extent possible, the benefits and risks associated with the technology.

The language used in the disclosures should be comprehensible to investors. This will help build trust and enable clients to understand the service and products that they are being offered and sold and allow them to make informed decisions.

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

The performance of AI and ML is largely dependent on data quality and lack of bias in processing. The existence of bias in the results of AI and ML algorithms may jeopardise both those firms that use them and their customers, due to the risks of discrimination or inadequate advice.

Firms should therefore ensure an adequate level of data quality, by checking the quality of the sources used, as well as the relevance and completeness of the data with regard to the underlying objectives of the algorithm. In particular, firms should ensure that the data set is representative of the target population so that they do not lead to exclusion phenomena.

Biases may exist both in the data collected and in the manner in which they are processed. They may be present directly in the variables used, for example, with variables considered discriminatory such as gender. They can be implicit due to the interaction of several variables that are not in themselves discriminatory. Firms should ensure that the data as well as the outputs of the AI and ML are analysed for the risk of discrimination.

Biases can be reinforced by AI and ML algorithms, resulting in consumer harm. For example, models based on behaviour history are performing sub-optimally for younger customers with a limited history, in which case other explanatory variables need to be found.

Firms should have appropriate processes and controls in place to identify and remove biases from data sets. Firms should consider having specific training courses to raise awareness amongst their data scientists (and/ or other relevant staff) of potential data biases.

Chapter 7 – Conclusion and Next Steps

The proposed guidance intends to address some of the potential risks surrounding the general use of AI and ML by market intermediaries and asset managers. If implemented, the proposed guidance should help ensure that firms have adequate control frameworks to appropriately use AI and ML.

Definition
Question 1: Do you agree with the proposed definition of AI and ML?

Risks and challenges
Question 2: Do you see any risks or challenges around AI and ML which are not mentioned in the report?

Guidance
Question 3: Do you agree that the guidance set out in Chapter 6 of the Consultation Report is appropriate to address the potential risks associated with the general use of AI and ML by market intermediaries and asset managers? If not, please provide details.

Disclosure of Information
Question 4: Do you disclose information to customers / clients on the use of AI and ML? If yes, please indicate what kind of information is disclosed.

Question 5: What factors do you need to take into account when considering the appropriate level of information that should be disclosed to clients (including prospective clients) and other relevant stakeholders around the firm's use of AI and ML algorithms?

Proportionality
Question 6: How do you consider and apply proportionality to your systems and controls over the use of AI and ML?

Annex 1

How regulators are addressing the challenges created by AI and ML

FSRA (Abu Dhabi Global Market)

The Financial Services Regulatory Authority (FSRA) issued detailed guidance in July 2019 setting out the regulatory framework for firms carrying out "digital investment management" (i.e., robo-advisory activities) in Abu Dhabi Global Market (ADGM).[12]

Such firms (termed "Digital Investment Managers") are required to apply for a licence from the FSRA to conduct one or more regulated activities which may include (depending on their business models):

- Advising on Investments or Credit
- Arranging Deals in Investments
- Managing Assets

Digital Investment Managers – being regulated entities – would therefore need to comply with the FSRA's regulatory requirements such as those pertaining to the conduct of business and prudential requirements. Such firms are also required to have robust frameworks and controls in place with respect to client suitability and disclosures, algorithm governance, and technology governance.

The limited human interaction between Digital Investment Managers and their clients necessitates careful consideration of how suitability assessments are performed. Digital Investment Managers typically rely heavily on an online questionnaire to collect the information needed to perform suitability assessments ("Risk Profile Questionnaire"). When designing a Risk Profile Questionnaire, the FSRA expects that Digital Investment Managers ensure that the information obtained to assess suitability is proportionate with the complexity and risk of the products being sold. In addition, the Risk Profile Questionnaire should have mechanisms for "knock out" questions that reject prospective clients whose investment horizon, liquidity needs or other circumstances are misaligned with the products offered through the platform.

A critical component of the digital investment management business model is the use of algorithms to automate the investment process. Accordingly, the FSRA sees a need to ensure that Digital Investment Managers have adequate algorithm and technology governance policies and processes in place to address the specific risks arising from such a technology-driven business model.

From a governance perspective, the FSRA expects that Digital Investment Managers will establish internal governance structures that enable its Board and Senior Management to have robust oversight and control over the design, performance, deployment and security of algorithms. The roles and responsibilities of all personnel who oversee the design, performance and integrity of algorithms must also be clearly defined.

[12] Abu Dhabi Global Market (ADGM) was established on 21 October 2015 as a broad-based international financial centre in the Emirate of Abu Dhabi, with its own civil and commercial laws independent from the rest of the United Arab Emirates.

In terms of the development and testing of the algorithm, Digital Investment Managers must maintain proper documentation explaining the decision tree or logic of the algorithm to ensure that the outcomes produced by the model are explainable, traceable and repeatable. The Digital Investment Manager must also ensure the relevance of any data or assumptions upon which the algorithm is based, and that the Risk Profile Questionnaire completed by clients takes into account potential behavioural biases that may lower the accuracy of client responses. The Digital Investment Manager must carry out sufficient testing to demonstrate that its algorithm meets these principles. Where appropriate (e.g., in the case of a complex algorithm), the FSRA may require a third-party audit to validate the performance outcomes of the algorithm as purported.

Additionally, given their heavy dependence on collecting and processing client data and the risks of cyberattacks to their automated and largely digital mode of operations, the FSRA requires Digital Investment Managers to implement robust data security policies and systems to ensure compliance with all relevant data protection regulations, including the ADGM's Data Protection Regulations.

FCA (United Kingdom)

It is essential that key oversight functions, including compliance and risk management, keep pace with technological advancements. In the absence of appropriate systems and controls, the increased speed and complexity of financial markets can turn otherwise manageable errors into extreme events with potentially wide-spread implications.

The Financial Conduct Authority (FCA) deems it good practice to review how trading algorithms are used; develop appropriate definitions; ensure all activities are captured; identify any changes to algorithms; and have a consistent methodology across the testing and deployment of AI and ML. Markets in Financial Instruments Directive (MiFID II) requires firms to develop processes to identify algorithmic trading across the business. These can be either investment decisions or execution algorithms, which can be combined into a single strategy. Firms are also required to have a clear methodology and audit trail across the business. Approval and sign-off processes should ensure a separation of validation and development a culture of collaboration and challenge and consistency of a firm's risk appetite. Whilst the algorithms are field-deployed, it is a requirement to maintain pre-trade and post-trade risk controls, real-time monitoring of algorithms in deployment, with the ability to kill an algorithm or a suite of algorithms centrally, a functionality commonly known as the kill-switch.

It is a best practice, but not a requirement, to have an independent committee to verify the completion of checks. However, under the SM&CR, a firm's governing body would be expected explicitly to approve the governance framework for algorithmic trading, and its management body should identify the relevant Senior Management Function(s) with responsibility for algorithmic trading.

Canada

With respect to the algorithmic trading method in Canada, under National Instrument *23-103 Electronic Trading and Direct Electronic Access to Marketplaces*[13] and the Investment Industry Regulatory Organization of Canada (IIROC) *Notice 12- 0364 – Guidance Respecting Electronic Trading*, firms are required to ensure that algorithmic trading systems used by participants to route orders to marketplace have been adequately tested prior to being "engaged", including the ability of the firm to immediately disengage the operation of the algorithmic trading system should the need arise. It is expected that firms ensure, at a minimum, that each algorithmic trading system has been tested under various market conditions to identify problematic outcomes related to the operation of the algorithmic trading system.

Also, firms should have built-in features or functionality that prevent (or provide real-time alerts) certain pre-programmed order or trade parameters from being exceeded (e.g. certain volume, order or price limits). Firms should also have an "override" functionality which either automatically "disengages" the operation of the algorithmic trading system or permits the firm to do so remotely. Additionally, as part of their supervisory policies and procedures, firms are required to maintain a written record with sufficient details to demonstrate the testing of the algorithmic trading system undertaken, including details of the testing conducted by the algorithmic trading system provider.

Further, at the national level, consumer protection and personal information protection laws are expected to apply to AI tools and services. Quebec AMF is of the view that it might become necessary to make new stakeholders accountable for their operations or to request certification/IT audit of the underlying systems. Except as expressly stated, existing laws and regulations administered by the Canadian Securities administrators (CSA) are "technology neutral". For example, as stated in the CSA Staff Notice *31-342 Guidance for Portfolio Managers Regarding Online Advice,*[14] the applicable rules are the same if the activities are conducted under the traditional model of interacting with clients face-to-face and if a portfolio manager uses an online platform.

On December 4, 2018, the Montréal Declaration for Responsible AI[15] was published, a paper drafted by leading AI institutes, AI researchers, and other stakeholders in Canada, aimed at helping guide individuals, organisations, companies and governments make responsible and ethical choices when building and utilising AI technology. The Montréal Declaration for Responsible AI provides effective supervisory and oversight principles for AI and ML development and deployment.

BaFin (Germany)

---

13 CSA National Instrument 23-103 *Electronic Trading and Direct Electronic Access to Marketplaces*, available at: https://lautorite.qc.ca/fileadmin/lautorite/reglementation/valeurs-mobilieres/23-103/2014-03-01/2014mars01-23-103-vofficielle-en.pdf

14 CSA Staff Notice 31-342 *Guidance for Portfolio Managers Regarding Online Advice*, September 2015, available at: https://lautorite.qc.ca/fileadmin/lautorite/reglementation/valeurs-mobilieres/0-avis-acvm-staff/2015/2015sept24-31-342-avis-acvm-en.pdf

15 Montréal Declaration Responsible AI, December 2018, available at: https://docs.wixstatic.com/ugd/ebc3a3_c5c1c196fc164756afb92466c081d7ae.pdf

BaFin published its study "*Big data meets artificial intelligence*" in 2018. The study was triggered by the fact that the entities BaFin supervises store and analyses data in very different ways, and that this impacts how financial services are being provided. Its objective is to provide an understanding of these trends as a basis for discussing the supervisory and regulatory implications. The report stresses consumer trust and data sovereignty as key success factors of BDAI applications and innovation and identifies the following supervisory and regulatory implications:

1. Supervision of firms

a) BDAI Governance

- When designing partially automated processes, it is important to ensure that they are embedded in an effective and appropriate manner. Ultimately, responsibility remains with the senior management of the firm. Appropriate documentation is required to ensure this.

- It is the responsibility of firms to guarantee the **explainability** (the ability to determine the main factors influencing a specific individual decision that has been reached by a system") and traceability of BDAI-based decisions. The study shows how new approaches could provide at least some insight into how models work and the reasons behind decisions, even in the case of highly complex models, thereby preventing models from being categorised purely as black boxes.

- Supervisory and regulatory authorities will therefore **not accept** any models presented as an unexplainable **black box**. In addition, a better understanding of models provides an opportunity to improve the analysis process – allowing, for instance, the responsible units in the supervised firm to identify overfitting and data bias.

- According to the study, it may be necessary to further develop established **governance concepts**, such as the principle of dual control, and to apply these to automated processes. For example, one could think about introducing special safeguards for certain particularly risky BDAI applications, safeguards already used in other technological applications (in aviation, speed is measured using several independent systems - the back-up would be an algorithm).

- Due to the complexity of the applications, also resulting from the type and scope of data used, consideration should be given as to whether the process results should also be **monitored** in the future, in addition to the **documentation requirements** described above. For example, this could be done by evaluating the results produced by an algorithm in a test scenario set by the supervisory authorities.

b) Fighting financial crime and preventing conduct violations:

- Exploiting BDAI potential in compliance processes - BDAI can improve the detection rate of anomalies and patterns, and thus increase the efficiency and effectiveness of compliance processes, such as money laundering detection or fraud prevention. In addition, BDAI can also play a role in monitoring employee conduct.

- Preventing criminals from turning to less advanced firms - Should the use of BDAI technologies result in far more efficient ways to detect money laundering, criminals could potentially turn to firms that are less advanced in this area. It is therefore necessary to monitor whether this will materialise.

- Defining supervisory requirements for the explainability and effectiveness of algorithms for detecting financial crimes - Supervisory and regulatory authorities need to discuss whether it is necessary to define specific minimum standards for the explainability/traceability and effectiveness of the methods used. The results of the algorithms would have to be traceable to the extent that the supervisory authorities can monitor them and that the competent authorities, e.g., law enforcement agencies, can make use of them. In addition to performing comparisons with specific minimum standards that may need to be developed, supervisory authorities could also determine the effectiveness of algorithms by benchmarking them against the algorithms of other providers.

c) Internal models subject to supervisory approval

- Possible uses of BDAI in **models requiring supervisory approval** - Supervised firms may also wish to use BDAI methods in models that are subject to approval by the supervisory authority, for example, the internal ratings-based approach for banks. This is likely to be the case if the use of BDAI leads to more precise results than the combination of methods and data currently used. In this way, BDAI could potentially be used to improve current methods, for instance, those used to optimise the assessment of individual counterparty default risks. In addition, a better portfolio view could possibly alter the weighting of individual risks as concentration risks could be addressed more appropriately, for instance, by analysing the degree of interconnectedness between borrowers.

- Actual effects of BDAI use on capital requirements are likely to be limited. It is worth noting that there are currently no BDAI applications observed or authorised in the algorithms of such models. It should also be noted that the authorised models in existence today are highly advanced, for instance, with regard to the scope of the data used, including external data. In the banking sector, it should also be borne in mind that the output floor approved as part of the finalisation of the Basel III reforms would additionally limit the actual effects on capital requirements.

- Defining prerequisites for **BDAI use in models** requiring **supervisory approval** -
  The use of BDAI models would always be subject to a corresponding approval on a case-by-case basis. Beyond the individual case, the question could be asked as to whether all BDAI methods are equally suited for use in models that require supervisory approval, or whether there are methods that should be ruled out per se.

  Furthermore, it is necessary to examine whether existing legal (minimum) standards for the data used and for BDAI model transparency are sufficient or whether additional requirements would be necessary.

d) Handling information security risks

- Information security risks increase as a result of greater use of BDAI - The growing complexity caused by BDAI presents new challenges in managing information security risks. Together with ever larger data volumes, the disaggregation of value chains supported by BDAI creates a **larger attack surface** while simultaneously reducing each individual provider's ability to control the data used and distributed.

- In addition, a new phenomenon has arisen: **attacks on certain BDAI algorithms** through data manipulation, e.g., in the form of adversarial or poisoning attacks. This leads to risks that are very different to those encountered in traditional online banking, resulting in a rise in both operational and reputational risks.

- BDAI and encryption systems for **mitigating information security risks** - Although the use of BDAI can result in an increase in information security risks, BDAI can also be used to mitigate such risks, for example, to analyse and detect danger. For instance, BDAI could be used to identify irregularities in user behaviour in online banking, which could indicate potential improper use. Certain encryption systems that allow BDAI methods to be used directly on encrypted data could also be used to reinforce resilience against such risks.

China

In order to effectively cope with the challenges brought to intermediaries by new technologies such as AI and ML, the CSRC issued the " *Measures for the Administration of Information Technology Management of Licensed Corporation*" in December 2018 (CSRC No. 152, hereinafter referred to as "Measures").

First, this promotes the integration of information technology governance and corporate governance, thus consolidating the foundation of information technology governance. The “Measures” requires intermediaries to clarify the IT management responsibilities of senior management. Also, it improves the distribution of power and responsibility through the establishment of CIO and information technology governance committees.

Second, the “Measures” requires intermediaries to ensure that the back-office management capabilities should be consistent with the complexity of front-office application, strengthening compliance and risk management. Taking into account the professionalism and complexity of AI and ML, the “Measures” requires the compliance management department and the risk management department to equip with information technology resources that are compatible with the scale and complexity of business activities, and establish corresponding review, monitoring and inspection mechanisms.

Third, the “Measures” requires the enhancement of data governance. Due to the importance of data for AI and ML, it encourages intermediaries to exploit full value of data and explore data applications in business operations, risk management and internal control. On the other hand, adhering to the principle of “Tech for Social Good”, it requires intermediaries to strengthen data life-cycle management, including data security management, log recording and prohibitions on over-collect and customer data abuse.

Fourth, it requires to strengthen the management of third party providers. Given the fact that a large number of AI and ML applications are provided by third party providers, the regulation

of third party providers is critical. Therefore, the "Measures" clarifies the responsibilities and behavioural boundaries between intermediaries and third-party provider and requires third party providers to file with CSRC if important information system services are provided to intermediaries.

Further improvements on regulatory measures based on CSRC's recent practices.

First, based on traditional risk management framework, regulators should require intermediaries to adopt technical methods to improve compliance and risk management capabilities, ensuring those are compatible with the technology applied in the industry.

Second, regulators should require intermediaries to carefully select service providers, and continue to pay attention to service concentration, in order to prevent convergence of trading strategies, triggering market sentiment resonance, and extreme circumstance in which large number of intermediaries are affected at the same time.

AMF (France)

In France, the Autorité de contrôle prudentiel et de résolution published, in conjunction with several other French financial regulators, including the Autorité des marchés financiers (AMF), a guidance[16] stating that the use of artificial intelligence for the allocation of assets or internal modelling of capital requirements may question the governance and risk management rules of the firm. Additionally, the reliability of algorithm, which first comes from data quality, is instrumental in the decision-making process. Some precautions should be taken, notably minimising the use of public external data, performing regular data quality checking and regular update of personal data with customers themselves.

To achieve these goals, several approaches may be considered, including:

- Use of experts to validate the relevance of the variables used, eliminate those that are unnecessary or a source of potential biases;
- Use of safer and more traditional parallel processes on part of the test data;
- Use of a standard dataset on the algorithms to regularly monitor both the relevance and non-discriminatory aspect of the algorithms; and
- Develop tools that would assess conceptual drift to control this specific risk of automatic learning.

Also, there should be conditions for monitoring these algorithms to demonstrate their explainability. The firms should be able to explain:

- Generally, the mechanisms and criteria followed by the algorithm during its analysis process;

[16] *Artificial intelligence: challenges for the financial sector*, December 2018, available at: https://acpr.banque-france.fr/sites/default/files/medias/documents/2018_12_20_intelligence_artificielle_en.pdf

- For a given action such as a decision taken, or an advice provided, the objective criteria and discriminant elements that conducted the algorithm, in the case examined, to effect one action or propose a solution rather than another

Coupled with the explainability test, a number of tests on datasets - independent from those used for algorithm learning - could be considered to assess the quality of the results and the methodology for such tests remains should be defined.

Furthermore, some firms may consider maintaining a human intervention to verify the consistency of the results of the algorithm. This may be the case in the sensitive areas of information and advice provided to clients or AML-CFT. There may also have some challenges relating to financial stability. For example, there are several algorithms coded with similar variables, which may cause the high frequency trading programmes to converge towards the same strategy. The resulting risk is to increase the pro-cyclicality and market volatility through simultaneous purchases and sales of large quantities. It would be therefore crucial to raise the bar of the risk analysis and management in such case to reduce more considerably the price differentials.

MAS (Singapore)

The Monetary Authority of Singapore (MAS) issued a paper[17] setting out principles to promote fairness, ethics, accountability and transparency in the use of artificial intelligence and data analytics in financial sector. This paper contains a set of generally accepted principles for the use of AI and data analytics in decision-making in the distribution of financial products and services, with a particular mention on building consumer confidence and trust in providing their personal data for AI, such as selecting appropriate organisational governance measures and adopting good data management practices. In developing this set of principles, MAS has worked closely with key industry stakeholders and sought feedback from a range of financial institutions and companies, who have offered diverse perspectives and inputs. The paper provides some guidance on the four following principles:

- Fairness (justifiability, accuracy and bias)
- Ethics
- Accountability (internal and external)
- Transparency

The paper also provides some illustration on actions to be taken by the firms to ensure they comply with the principles. For example, the frequency of validation or tests of algorithms depends on the materiality of the decision and the complexity of the model. For particularly material and complex deep learning models, a firm requires more frequent validation and testing. Conversely, for less material and more straightforward models, quarterly, annual or even less frequent validations and tests suffice.

AFM (Netherlands)

[17] Principles to Promote FEAT in the Use of AI and Data Analytics in Singapore's Financial Sector, November 2018, available at: https://www.mas.gov.sg/-/media/MAS/News-and-Publications/Monographs-and-Information-Papers/FEAT-Principles-Updated-7-Feb-19.pdf

The AFM believes it is vital to have key functions within the firm to safeguard the safety and use of AI and ML techniques within the organisation. The risk and compliance functions play an important role and should preferably be involved during the initial development process. Senior management can set the conditions for AI and ML techniques and take responsibility for techniques deployed in the firm. A clear governance framework will reflect the firms control on its AI and ML techniques and on the data it feeds into the models.

Concerning AI and ML techniques and the data itself the AFM believes it to be good practice to check the quality and accuracy of the input data; make conscious decisions on the specific AI and ML technique deployed; determine the level of explainability suitable to become and stay in control; be aware of the effect of using third party providers; and have the skills in place to test, monitor and challenge both internally and externally developed datasets and AI and ML techniques. Clear documentation of the considerations and decisions made in the process help to understand why firms chose a specific path.

AI and ML techniques also create new opportunities for firms to further strengthen and safeguard the customer interest.

Securities and Exchange Commission (United States)

In October 2018, the Securities and Exchange Commission (SEC) launched the agency's Strategic Hub for Innovation and Financial Technology (FinHub). The FinHub is a resource for public engagement on the SEC's FinTech-related issues and initiatives, such as artificial intelligence/machine learning, distributed ledger technology (including digital assets), automated investment advice, and digital marketplace financing.[18] The FinHub represents the latest step in the SEC's ongoing engagement related to artificial intelligence and machine learning.

With respect to robo-advisers, the staff of the Division of Investment Management at the SEC issued a guidance update in February 2017 that addressed certain unique considerations for robo-advisers to keep in mind as they seek to meet their fiduciary and substantive regulatory obligations under the Investment Advisers Act of 1940 (e.g., robo-advisers' unique business models, such as the level of overall human involvement, potentially limited face-to-face interaction with clients, and the use of algorithms).[19] Among other things, the guidance identified the following considerations: (i) the substance and presentation of disclosures to clients about the robo-adviser and the investment advisory services it offers; (ii) the obligation to obtain information from clients to support the robo-adviser's duty to provide suitable advice; and (iii) the adoption and implementation of effective compliance programs reasonably designed to address particular concerns relevant to automated advice. The staff also noted that there may be a variety of means for a robo-adviser to meet its obligations to its clients under the Advisers Act, and that not all of the issues addressed in its guidance will be applicable to every robo-adviser.

Financial Industry Regulatory Authority (United States)

---

18 For a list of relevant speeches and statements on AI/ML go to https://www.sec.gov/finhub .

19 *Robo-Advisers*, IM Guidance Update No. 2017-02, February 2017, available at: https://www.sec.gov/investment/im-guidance-2017-02.pdf

The Financial Industry Regulatory Authority (FINRA) has solicited broker-dealers' views on how FINRA can support FinTech innovation and helped educate firms on new regulatory issues. On July 30, 2018, FINRA issued a Special Notice on Financial Technology Innovation. The Special Notice requested comment on financial technology innovation in the broker-dealer industry. In the Special Notice, FINRA specifically requested comment regarding supervision in the context of artificial intelligence. FINRA noted that as broker-dealers increase their use of artificial intelligence, including chat bot-based services, they are grappling with how these tools may fit in the current regulatory framework. FINRA pointed to examples that included its supervision rule and previous guidance. FINRA Rule 3110 (Supervision) requires a broker-dealer to establish and maintain a system to supervise the activities of its associated persons that is reasonably designed to achieve compliance with the applicable securities laws and regulations and FINRA rules. In addition, FINRA restated that: "[A]s the use of algorithmic strategies has increased, the potential of such strategies to adversely impact market and firm stability has likewise grown. When assessing the risk that the use of algorithmic strategies creates, firms should undertake a holistic review of their trading activity and consider implementing a cross-disciplinary committee to assess and react to the evolving risks associated with algorithmic strategies." In the Special Notice, FINRA sought to better understand the challenges or issues that broker-dealers may face in deploying artificial intelligence tools and requested comment on any measures that FINRA could take to clarify or adapt its rules and processes in light of the evolving uses of such tools. Relatedly, FINRA issued Regulatory Notice 15-09 to address supervision and control practices for firms that engage in algorithmic trading strategies.

With respect to educating firms on new regulatory issues, FINRA issued a report on Technology Based Innovations for Regulatory Compliance ("RegTech") in the Securities Industry in September 2018. The report was a result of a focused review to learn more about the emergence and adoption of RegTech tools within the securities industry. The report discussed the use of machine learning with respect to regulatory intelligence programs, surveillance and monitoring, and investor risk assessments. FINRA encouraged comments on the report, including areas where guidance or modifications to FINRA rules may be desired to support adoption of RegTech tools while maintaining investor protection and market integrity. In addition to the report, FINRA hosted the 2019 FINRA RegTech Conference. The conference brought together regulators, thought leaders and industry practitioners to learn more about the use of RegTech tools, and related opportunities and challenges. Panelists discussed key innovative technologies that are transforming compliance functions including artificial intelligence.

FINRA also included regulatory technology as a highlighted item in the 2019 FINRA Risk Monitoring and Examination Priorities Letter. FINRA recognised that broker-dealers are using a variety of innovative RegTech tools to make their compliance efforts more efficient, effective and risk-based. As stated in the letter, "FINRA will engage with firms to understand how they are using such tools and addressing related risks, challenges or regulatory concerns, including those relating to supervision and governance systems, third-party vendor management, safeguarding customer data and cybersecurity."

CSSF (Luxembourg)

In December 2018, the Commission de Surveillance du Sector Financier (CSSF) issued a white paper on Artificial Intelligence describing opportunities, risks and recommendations for the

financial sector.[20] The paper has no binding value vis-à-vis the supervised institutions but it includes practical guidance regarding the risks associated with the use of AI and ML and related recommendations to take into account.

Among the risks covered in the paper, there are data-related risks (e.g., data quality). Given the importance of data in AI and ML, having a solid data governance in place (e.g., setting clear roles and responsibilities for data ownership and defining processes to identify and fix data quality issues) is paramount to ensure that the data is of good quality. In case firms use external data sources, this requires having first verified their reliability (e.g., via due diligence of the external data service provider) and ensuring the data sources are adequate for their intended use. In addition, they should carefully review the integration of the AI and ML into the business process and establish controls performed by humans whenever there is a critical decision step.

In addition, the integration of the AI and ML into the business process should include controls performed by humans whenever there is a critical decision step.

An important section of the paper focuses on ethical aspects such as bias and discrimination, data privacy, accountability, explainability and auditability. The importance of these aspects is proportional to the potential impact that the AI and ML solution might have on the customer. The paper provides an overview of some existing techniques to help detect/prevent bias and to improve explainability of black box solutions.

Institutions should assume clear responsibility and accountability for the actions and decisions taken by automated AI and ML systems and processes. This responsibility cannot be delegated to a machine. Ultimately, responsibility should rely on the senior management of the institution which integrates the AI and ML logic into its business processes. When using third-party solutions, clear liability provisions may be defined at contractual level.

Another important aspect mentioned in the paper concerns the monitoring and update of the AI and ML solution, to promptly detect performance deviations or discriminative results and ensure that results remain accurate over time.

The lack of adequate internal skills to maintain the AI solution may lead to difficulties in updating the model or over-reliance on external parties. Therefore, firms should ensure a sufficient level of internal AI skills to understand, develop and supervise the solution, and should consider setting up multidisciplinary teams involving data scientists, IT, risk and compliance staff. The paper also describes the main categories of security attacks affecting AI and ML solutions and related defence techniques, recommending institutions to perform technical watches over these continuously evolving topics, in order to remain up-to-date from a technical perspective.

Finally, another vigilance point concerns systemic risks, which may be caused, for example, by the use of a common trading algorithm by many institutions, especially if the algorithm contains errors. Therefore, there should be a monitoring of potential systemic effects to issue warnings when appropriate.

---

[20] Artificial Intelligence: opportunities, risks and recommendations for the financial sector, December 2018, available at: https://www.cssf.lu/fileadmin/files/Publications/Rapports_ponctuels/CSSF_White_Paper_Artificial_Intelligence_201218.pdf

Annex 2

Guidance published by supranational bodies

Financial Stability Board (FSB)

On November 2017, the FSB published a report considering the financial stability implications of the growing use of AI and ML in financial services.[21]

The report outlines that financial institutions are increasingly using AI and ML in a range of applications across the financial system including front office client interactions, risk management, trade execution optimisation, as well as regulatory compliance and fraud detection.

The FSB's analysis reveals a number of potential benefits and risks for financial stability that should be monitored as the technology is adopted and as more data becomes available. They are:

- The more efficient processing of information, for example in credit decisions, financial markets, insurance contracts and customer interactions, may contribute to a more efficient financial system. The applications of AI and ML by regulators and supervisors can help improve regulatory compliance and increase supervisory effectiveness.

- Applications of AI and ML could result in new and unexpected forms of interconnectedness between financial markets and institutions, for instance based on the use by various institutions of previously unrelated data sources.

- Network effects and scalability of new technologies may give rise to third-party dependencies. This could in turn lead to the emergence of new systemically important players that could fall outside the regulatory perimeter.

- The lack of interpretability or auditability of AI and ML methods could become a macro-level risk. Similarly, a widespread use of opaque models may result in unintended consequences.

- As with any new product or service, it will be important to assess uses of AI and ML in view of their risks, including adherence to relevant protocols on data privacy, conduct risks, and cybersecurity. Adequate testing and 'training' of tools with unbiased data and feedback mechanisms is important to ensure applications do what they are intended to do.

[21] Artificial intelligence and machine learning in financial services, Market developments and financial stability implications, November 2017, available at: https://www.fsb.org/wp-content/uploads/P011117.pdf

International Monetary Fund (IMF)

The IMF published "Fintech and Financial Services" in June 2017,[22] sets an economic framework for considering fintech channels (including AI and ML) that might provide solutions that respond to consumer needs for trust, security, privacy, and better services, change the competitive landscape and affect regulation. It highlights numerous potential opportunities for AI solutions within financial services including for example tracking regulatory changes and interpreting new regulations.

Additionally, the IMF report highlighted areas of its own work that could potentially utilise AI and ML once the issues of provenance of data, and data privacy have been addressed before basing analysis or grounding policy advice on AI and ML outputs.

Institute of Internal Auditors (IIA)

The IIA published guidance[23] introducing an AI auditing framework to help organisations evaluate, understand, and communicate the degree to which AI negatively or positively affects the organisation's ability to create value in the short, medium, or long term.

Information Systems Audit and Controls Association (ISACA)

ISACA published a white paper[24] discussing the information technology that auditors need to know when auditing AI reliant businesses in the absence of a mature auditing framework and AI-specific regulations or standards.

European Commission

The European Commission released draft ethics guidelines for AI in April 2019.[25] The paper highlights the importance of trust in ensuring a human-centric approach to AI and outlines three components that should be met throughout the AI system's life:

- It should be lawful, complying with all applicable laws and regulations;
- It should be ethical, ensuring adherence to ethical principles and values; and
  it should be robust, both at a technical and social level.

Furthermore, the paper provides guidance on how such principles can be operationalised in socio-technical systems. The guidance is being piloted by stakeholders and feedback will be presented to the European Commission in early 2020.

---

22 *Fintech and Financial Services*, June 2017, available at https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2017/06/16/Fintech-and-Financial-Services-Initial-Considerations-44985

23 *Artificial Intelligence, Internal Audit's Role, and Introducing a New Framework*, 2017, available at: Artificial Intelligence, Internal Audit's Role, and Introducing a New Framework

24 *Auditing Artificial Intelligence*, 2018, available at: Auditing Artificial Intelligence

25 *Ethics Guidelines For Trustworthy AI*, April 2019, available at: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top

The International Technology Law Association (ITechLaw)

ITechLaw[26] published a paper in May 2019[27] providing a principled framework that encourage the responsible development, deployment, and use of artificial intelligence. The principles relate to:

- Ethical Purpose and Societal Benefit
- Accountability
- Transparency and explainability
- Fairness and Non-discrimination
- Safety and Reliability
- Open Date and Fair competition
- Privacy
- AI Rights and Intellectual Property

Organisation for Economic Co-operation and Development (OECD) and the G20

On May 2019, OECD adopted principles along with recommendation on AI,[28] produced by a group of leading international experts. The OECD AI principles are the first such principles signed up by governments.

The OECD's five principles are:

- ***Inclusive growth, sustainable development and well-being*:** stakeholders should pursue beneficial outcomes, such as augmenting human capabilities and enhancing creativity, advancing inclusion of underrepresented populations, reducing economic, social, gender and other inequalities, and protecting natural environments.

- ***Human-centred values and fairness*:** AI actors should respect the applicable laws, human rights and democratic values, such as freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, fairness, social justice; as well as internationally recognised labour rights.

- ***Transparency and explainability*:** AI actors should disclose meaningful information to foster a general understanding of AI systems, make stakeholders aware of their interactions with AI systems, including in the workplace, enable those affected by an AI system to understand the outcome and those adversely affected by an AI system to challenge the outcome

- ***Robustness, security and safety*:** AI systems should be robust, safe and secure throughout their entire lifecycles. AI actors should ensure traceability of datasets, processes and decisions making during the AI system lifecycle to enable AI actors to

[26] A multi-disciplinary group of 54 technology law experts, researchers and industry representatives from 16 countries.

[27] Responsible AI: A Global Policy Framework, May 2019, available at: https://www.itechlaw.org/sites/default/files/ResponsibleAI_Principles.pdf

[28] *OECD Principles on AI*, May 2019, available at: http://www.oecd.org/going-digital/ai/principles/

analyse AI systems' outcomes. AI actors should also apply a systematic risk management approach throughout AI systems' lifecycles.

- ***Accountability*:** AI actors should be accountable for the proper functioning of AI systems.

The OECD makes five recommendations:

- Invest in AI research and development;
- Foster a digital ecosystem for AI;
- Shape an enabling policy environment for AI;
- Build human capacity and preparing for labour market transformation; and
- Encourage international co-operation for trustworthy AI.

The G20 endorsed these OECD AI principles and recommendation in June 2019.