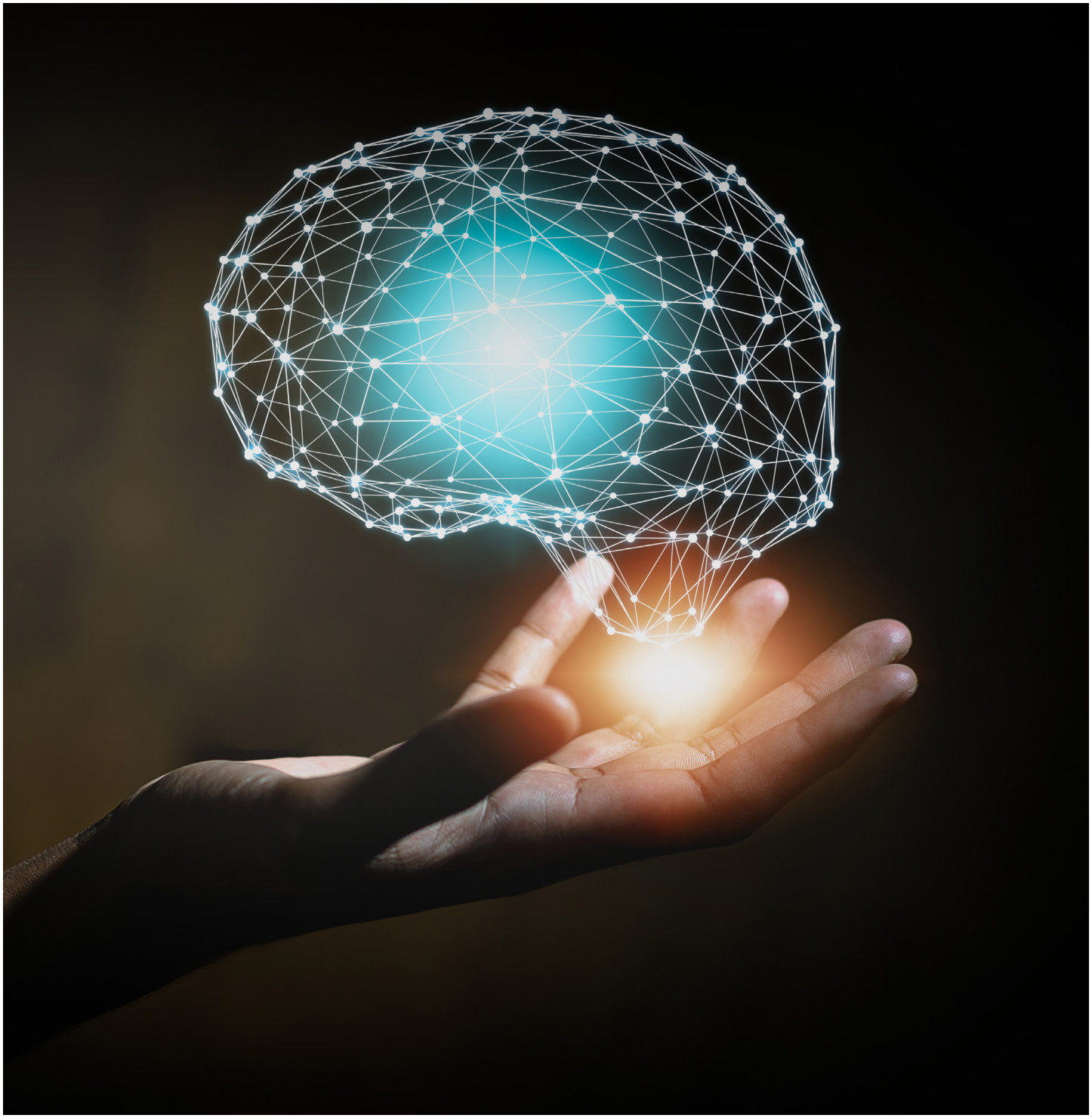
Enabling an Efficient Regulatory Environment for AI



JUNE 2021

asifma
Growing Asia's Markets

Disclaimer

The information and opinion commentary in this ASIFMA – Enabling an Efficient Regulatory Environment for AI (Paper) was prepared by the Asia Securities Industry and Financial Markets Association (ASIFMA) to reflect the views of our members. ASIFMA believes that the information in the Paper, which has been obtained from multiple sources believed to be reliable, is reliable as of the date of publication. As estimates by individual sources may differ from one another, estimates for similar types of data could vary within the Paper. In no event, however, does ASIFMA make any representation as to the accuracy or completeness of such information. ASIFMA has no obligation to update, modify or amend the information in this Paper or to otherwise notify readers if any information in the Paper becomes outdated or inaccurate. ASIFMA will make every effort to include updated information as it becomes available and in subsequent Papers.

asifma
Growing Asia's Markets

ASIFMA is an independent, regional trade association with over 140 member firms comprising a diverse range of leading financial institutions from both the buy and sell side including banks, asset managers, law firms and market infrastructure service providers.

Together, we harness the shared interests of the financial industry to promote the development of liquid, deep and broad capital markets in Asia. ASIFMA advocates stable, innovative and competitive Asian capital markets that are necessary to support the region's economic growth. We drive consensus, advocate solutions and effect change around key issues through the collective strength and clarity of one industry voice. Our many initiatives include consultations with regulators and exchanges, development of uniform industry standards, advocacy for enhanced markets through policy papers, and lowering the cost of doing business in the region. Through the GFMA alliance with SIFMA in the US and AFME in Europe, ASIFMA also provides insights on global best practices and standards to benefit the region.

www.asifma.org

1 Executive Summary

This section provides a summary of the report.

Artificial Intelligence (AI) is now seen as one of the megatrends in the financial industry. Many financial institutions have been adopting and introducing AI into their daily operations across a wide range of use cases including predictive data analytics, liquidity risk analysis, sentiment analysis, securities research, stock selection, voice-to-text natural language processing, smart matching of trades, market abuse and financial crime surveillance, credit scoring, marketing etc. Various regulatory agencies across jurisdictions in Asia-Pacific (APAC) have begun introducing guidelines relating to the use of AI. Because of these individual approaches, there is a risk of regulatory fragmentation and duplication of well-established regulation and standards that the financial industry is already subject to, which might stifle innovation and increase regulatory compliance risks. Individual approaches to AI among regulators and overly prescriptive rules could cause any rules or guidance to fall out of step with rapid technological developments, undermining their effectiveness, while also unnecessarily raising compliance costs and potentially hampering innovation.

Whilst there is no globally accepted definition of AI, within this paper, to understand the term "artificial intelligence" we refer to "systems that act in the physical or digital world by perceiving their environment through data acquisition, interpreting the collected data, reasoning on the knowledge, or processing the information derived from this data and identifying the best action(s) to take to achieve the given goal. AI systems adapt themselves or their own algorithms by analysing how the environment is affected by previous actions, knowledge or data."[1] In this paper, when referencing AI, this also covers Machine Learning (ML) as ML is a subset of AI.

Financial services is a highly regulated sector and we found in our research that existing regulations largely address and mitigate the key risks which might be caused or increased by the use of AI. These include rules in respect of outsourcing, technology risk management, conduct, cybersecurity, duty to clients, internal governance, and model risk management, in addition to sector-agnostic requirements around data privacy and data protection and established internal risk management and governance frameworks. Our research studying how existing regulations and guidelines govern risks relevant to the use of AI concludes that existing rules and guidelines largely address AI-related risks.

We recommend that regulators take a principles- and risk-based approach to AI, giving financial institutions flexibility in how best to operationalise the principles in relation to their AI adoption, depending on the financial institution's setup, framework, and the materiality of the AI use case. We encourage regulators to support the global development of AI within capital markets and avoid fragmentation and overregulation, which could slow down its adoption and development.

To that end, in Section 2 of this paper we propose a set of regulatory principles for AI which we believe will form the basis for an efficient regulatory environment whilst at the same time supporting customer and investor protection, market integrity and financial and systemic stability.

[1] AFME (2020): https://www.afme.eu/Portals/0/DispatchFeaturedImages/20200612%20AFME%20EC%20AI%20CP%20Response%20-%20Final_.pdf Accessed on 30 October 2020

Specifically, we recommend that regulators should:

- **Principle 1**
  Support public-private collaboration
- **Principle 2**
  Allow financial institutions to take a risk-based approach to AI, taking materiality of the use case and stakeholders into account
- **Principle 3**
  Take a technology-agnostic approach to regulation
- **Principle 4**
  Leverage existing regulatory frameworks
- **Principle 5**
  Strive for regional and international harmonisation
- **Principle 6**
  Promote and facilitate cross-border data flow
- **Principle 7**
  Engage with the industry on areas that need further discussion

We have described in Section 3 the existing general regulations and guidelines across a number of key focus areas which relate to the key risks AI poses in financial services, based on our working group's assessment:

- **Fairness**
- **Governance and Accountability**
- **Transparency**
- **Data Quality**
- **Data Protection**
- **Model Governance**
- **Resilience**
- **Cybersecurity**
- **Third Party Risk Management**
- **Expertise**

For each of these focus areas, we assessed how existing general regulations and guidelines pertaining to financial institutions address these risks, and whether there are any gaps.

Section 4 provides an overview of AI-specific guidelines issued to date. We have assessed these AI-specific guidelines against our focus areas.

In the Annex, we have reviewed the regulations relevant to AI in a number of APAC jurisdictions (Hong Kong, Singapore, China, India, South Korea, Japan, Thailand, Australia) and compared them with those in a number of leading financial services jurisdictions (European Union (EU), United States of America (US) and United Kingdom (UK)), including both pre-existing regulation in the areas of technology risk management, cybersecurity, data privacy, as well as an overview of more recently-issued AI-specific guidelines.

This Paper has been written by the ASIFMA AI task force, which consists of banks, asset managers, professional firms and market infrastructure providers. A special thanks goes to the following firms who were instrumental in supporting the group with the desktop research and mapping: Bae, Kim and Lee, Clifford Chance, EY, KPMG, King & Wood Mallesons and PricewaterhouseCoopers. In addition to the desktop research, ASIFMA engaged with the wider industry through a survey and follow-up interviews. The survey was distributed to ASIFMA members, as well as members of the Association of Banks in Singapore (ABS), Australian Financial Markets Association (AFMA), Alternative Investment Management Association (AIMA) and The Hong Kong Association of Banks (HKAB).

We look forward to engaging with regulators and other stakeholders on our suggested principles and key findings to support an enabling regulatory framework for AI in APAC and beyond.

2

ASIFMA Principles For AI Regulation

This section proposes seven principles that regulators should consider when taking policy actions in relation to AI.

AI has the potential to positively impact the financial markets industry on a global scale. The use of AI by financial institutions creates significant efficiencies and benefits for both financial institutions and investors such as increased transaction execution speed and lower costs of investment services [2]. As capital markets participants are increasingly adopting AI for a variety of use cases, regulators across the APAC region and globally are looking at their existing and future policy approaches.

We recognise that notwithstanding the benefits, there are also a number of potential risks associated with the development and use of AI. However, as we will show further in the paper, many of these risks are not new to the industry or specific to AI (e.g. governance, resilience, cybersecurity and data privacy) and should have already been sufficiently embedded in firms' existing risk management frameworks and addressed through existing regulations, covering numerous aspects of the business. These existing requirements apply to financial institutions regardless of whether the relevant processes use AI.

As an overarching recommendation, we recommend that regulators adopt a technology-agnostic, risk-based and principle-based approach to regulating AI. We recommend that financial institutions be regulated for these risks in relation to their business activities, irrespective of the technology used, unless that is attuned to real need. Additionally, it is important to ensure parity in expectations between non-AI and AI systems. The operational benchmark for AI systems should focus on the performance of comparable current processes (if existing) or an available human-powered alternative.

At the same time, there are a number of areas including transparency, explainability and fairness, that have special relevance to the use of AI that require further analysis and consideration. We encourage regulators to work collaboratively with the industry to address these areas whilst balancing market integrity, financial and systemic stability and customer protection on the one hand and innovation on the other hand. To that end, we welcome the recent effort by the International Organization of Securities Commissions (IOSCO) to seek public consultation on "Guidance on regulating the use of AI and ML by market intermediaries and asset managers".

We believe that AI can be used as a force for good in financial services when properly deployed as it can drive financial inclusion through innovation, lower the cost of financial services, tailor products to even better suit customer needs and profiles, reduce unlawful discrimination, increase efficiencies and improve risk management and financial crime and fraud prevention. Responsible and ethical use of AI and good AI governance is significant to the financial services industry. It is in the sector's best interest to work closely with the regulators, many of whom have already recognised the advantages that AI can bring to financial markets. We suggest that regulators take a supportive and encouraging stance towards the development and use of AI by financial institutions with the right guardrails and in an internationally coordinated fashion.

[2] IOSCO (2020): CR02/2020 The use of artificial intelligence and machine learning by market intermediaries and asset managers (iosco.org)

To that end, we hope that regulators consider the principles below when taking any policy actions in relation to AI:

Principle 1

Regulators should support public-private collaboration

Regulators should recognise that they and the industry continue to learn about the use and risks of AI. It is important, through public-private dialogue, for regulators to consider any concerns of financial institutions when considering issuing any rules or guidelines around AI. Input from skilled practitioners is also important to identify what is and what is not (yet) feasible and where unanticipated issues may arise. We believe regulators will also benefit from this public-private dialogue in terms of enhancing their own knowledge and skills to understand and regulate AI. Such collaborative public-private dialogue will be crucial to avoiding overregulation, maximising practicality and fully leveraging the benefits of AI.

AI, its use and regulation are quickly evolving and we therefore also encourage close collaboration between the regulators and industry through public-private partnerships and collaborative initiatives and allowing controlled experimentation in regulatory sandboxes at the option of the financial institutions. Regulatory sandboxes allow financial institutions or fintech firms to test their products, services or solutions often under a more relaxed regulatory environment but within well-defined limits and duration agreed with regulators. The aim is to support and facilitate the introduction of new innovative technologies and business models, and explore questions about the current regulatory framework or requirements for firms. An increasing number of jurisdictions, including but not limited to Hong Kong, Singapore, India, Korea, and the UK, have introduced regulatory sandboxes. Such controlled experimentation on the use and management of the risks of AI will help ensure the right balance is struck between risk mitigation and innovation to realise AI's full benefits.

Public sector-led initiatives on new technologies have proven successful, bringing together policymakers, regulators, academia, technology developers and industry participants for education and exploration of new opportunities. The Veritas consortium in Singapore is a good example of such public-private partnership. The Veritas consortium is a Monetary Authority of Singapore (MAS)-led framework to provide financial institutions with a verifiable way to incorporate MAS' Fairness, Ethics, Accountability and Transparency (FEAT) principles into their Artificial Intelligence for Data Analytics (AIDA) implementation. It will comprise an open source tool that will help financial institutions to assess their AIDA solutions against the FEAT principles (e.g. Veritas issued early 2021 a fairness assessment methodology for credit risk scoring and customer marketing[3]).Another good example is the Bank of England (BOE) and the Financial Conduct Authority's (FCA) Artificial Intelligence Public Private Forum[4] which was launched in October 2020. In Japan, the New Energy and Industrial Technology Development Organisation (NEDO) established by and under the Ministry of Economy, Trade and Industry (METI) has invited applicant vendors to conduct proof of concepts around the possible use of AI in anti-money laundering (AML)/ combating the financing of terrorism practices.

[3] MAS (2021): https://www.mas.gov.sg/news/media-releases/2021/veritas-initiative-addresses-implementation-challenges
[4] BOE (2020): https://www.bankofengland.co.uk/events/2020/october/fintech-ai-public-private-forum

Principle 2

Regulators should allow financial institutions to take a risk-based approach to AI, taking materiality of the use case and stakeholders into account

Regulators should ensure that any regulations applied to AI are applied on a basis that is risk-based, i.e., based on the materiality of the use case and the impact on client outcomes and markets. Proportionality should be a key principle for regulators in their oversight of AI use, and for financial institutions in their development and deployment of AI.

AI can potentially be used in a whole range of functions across financial markets, to augment existing activities, to replace them, or to perform complex and intensive tasks that were not previously feasible. Each use case will have its own risk profile and key stakeholders. Financial institutions should be allowed to adopt a risk-based approach that is proportionate to the risk of the particular use case and its potential impact on stakeholders, and measured against the performance of comparable current processes (if existing) or an available human-powered alternative.

As part of financial institutions' risk assessments, the following factors may be considered:

- The materiality of the activity and the extent to which AI is applied;
- The complexity of the AI model;
- The expected harm from the use of AI if it were to malfunction or produce an incorrect result;
- How the risks that already exist prior to using AI may be reduced or enhanced once AI is applied;
- The regulatory obligations, industry standards and internal policies or procedures to which the activity or technology is subject;
- The types of stakeholders who will be involved or affected, and any impact on clients, markets or counterparts; and
- The nature and sensitivity of data being used.

It is very important to ensure that regulators are adopting a differentiated approach depending on who are the stakeholders that are impacted by the use of AI. For instance, the level of controls should be higher when retail investors are directly impacted versus when the innovation is purely used within the financial institution with no impact on retail investors, markets or systemic risk. For example, those who develop AI for algorithmic trading and trade execution will need to consider the potential risks to clients and markets, ensuring that there are sufficient controls in place and that testing is undertaken on the application's performance under stressed market conditions. On the other hand, an AI application designed to manage a non-critical operational process (e.g. internal automation, removal of duplication of entries across multiple databases, marketing, natural language processing applications to extract information from documents, auto-routing of customer queries) may have a lower risk profile and may call for a different level of controls around assumptions and testing [5] . There could even be different levels of materiality for the same use case depending on the subject matter e.g. critical data that would directly affect markets or clients, versus internal non-sensitive data that may optimise internal workload allocation. A one-size-fits-all approach should therefore be avoided as it risks over-regulating inconsequential low-risk use cases, which would hinder innovation and slow down adoption of a technology that could have a beneficial impact. This need for proportionality was flagged by IOSCO in its June 2020 consultation paper on AI and ML. In the EU, the European Commission Proposal for a Regulation laying down harmonised rules on Artificial Intelligence[6] intends to prohibit certain AI practices which are considered to 'materially distort a person's behaviour in a manner that causes psychological or physical harm', or that 'exploit vulnerabilities of a specific group of persons'. Certain

[5] AFME (2019): https://www.afme.eu/Portals/0/DispatchFeaturedImages/100919%20AI%20transparency%20paper%20FINAL.pdf
[6] EC (2021): Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) | Shaping Europe's digital future (europa.eu)

activities are also classified as 'high risk', which the EU requirements prescribe for additional risk management systems and oversight to govern the use of AI deemed as 'high risk'. While ASIFMA does not intend to propose a ratings system to risk rate specific AI use-cases, the EU approach reinforces our suggestion for a risk-based approach to regulating AI.

Principle 3

Regulators should take a technology-agnostic approach to regulation

We recommend that regulators continue to adopt a technology-agnostic approach and that the focus should be on activities and outcomes, rather than seeking to regulate any particular technology use. Regulation should be technology-neutral and apply the principle of same activity, same risk, same regulation. We believe that technology-neutral requirements will allow any framework to remain dynamic and future-proof. It will ensure that regulation is able to keep pace with new technological developments, encourage innovation and not place unnecessary obstacles on the industry's use of the technology.

Regulating the risk / activity regardless of the entity type, instead of the technology will also ensure a level playing field, avoid overlapping and potentially inconsistent requirements and opportunities for regulatory arbitrage and support customer protection.

Principle 4

Regulators should leverage existing regulatory frameworks

As Section 3 shows, the financial industry is already highly regulated in relation to governance, risk management, accountability, cybersecurity, conduct, algorithm testing resilience, outsourcing, third party risk management and data privacy and have in place oversight structures dealing with the use of technology. Regulators should recognise this. Before considering new regulation, regulators should start by determining if these existing regulations already adequately address the identified risks, or if they need to be adapted to cover the risks AI presents or if greater clarity is needed on the applicability of existing requirements to AI – for example, through industry workshops, guidance notes and FAQs.

Many existing categories of regulation are largely technology-neutral, applying equally to manual processes and to sophisticated technology such as AI systems, or focussing on the deployment of technology generally (without necessarily differentiating between systems). Many of these requirements already drive the way that firms are developing and adopting AI. Our research in Section 3 shows that AI governance should and can fit within many existing risk management frameworks and that those existing frameworks can be leveraged to address any specific risks identified in relation to AI.

We welcome the March 2021 US joint agencies "Request for information and comment on the use of AI by financial institutions" [7]. This consultation does not presuppose the need for new AI-specific regulations, but provides an appendix of existing laws, regulations and guidance that may be relevant to AI, and limits the request for feedback to just clarification on compliance with applicable requirements.

[7] Federal Register (2021): Federal Register : Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning

We also recommend that regulators should – where necessary - take a progressive, incremental approach to regulation and should first consider whether it is possible to take non-regulatory action such as by the issuance of non-binding guidance, and interpretations on how financial institutions can adapt existing requirements to AI (for example, how financial institutions can establish a robust governance structure based on existing firm-wide risk management frameworks), before issuing any binding regulations, as necessary. The progressive nature of regulation should also extend to the content itself – commencing with principles and only becoming more specific if the risks and other circumstances warrant this. It should also extend to the manner of any enforcement, particularly as firms adapt.

We believe the introduction of detailed regulation of AI activities in banking and capital markets is not the right approach. Reactive and overly prescriptive responses tend to have unintended consequences and will hamper innovation in a space that is rapidly evolving and where firms are at early stages of implementation. Regulatory sandboxes and other public-private collaboration (per our Principle 1) would better serve as an avenue for some financial institutions who require a guided implementation and regulatory verification of their innovative AI applications.

Principle 5

Regulators should strive for regional and international harmonisation

As Section 4 shows, regulatory fragmentation occurs in APAC given the number of jurisdictions, their diversity and the lack of any regional supranational harmonising effort in financial regulation. Fragmentation of approaches to AI between jurisdictions adds additional cost, complexity and risks for financial institutions, which limits the potential benefits for both financial institutions and their clients. It also creates risk as it prevents firms from operating their business consistently and requires firms to create multiple approaches to execute the same services. As such, we recommend that regulators should seek to harmonise definitions and approaches across jurisdictions where possible. Regional and international consistency and compatibility is needed for global financial institutions to be able to use AI on a cross-border level. International regulatory forums and networks provide important opportunities to share best practices and identify specific cross border issues. The Global Financial Innovation Network (GFIN), Bank of International Settlements Innovation Hub, IOSCO FinTech Network and the Organisation for Economic Co-operation and Development (OECD) Global Partnership on AI[8] are some examples.

Principle 6

Regulators should promote and facilitate cross-border data flows

Regulators should support cross-border data flow with appropriate controls and continue to identify, avoid or eliminate forced data localisation requirements which impede the development of technology-driven economic growth including innovations such as AI.

Laws and regulations that restrict the cross-border flow of data - including data localisation or residency requirements - significantly hinder the development and use of AI. The ability to access and process large datasets to feed into AI models, is crucial for innovations in AI including ensuring development of holistic, quality datasets. This means that it is imperative for data to flow across borders in order to allow for AI innovations to flourish in country. We recommend that regulators adopt policies that encourage the free-flow of data across borders. In contrast, data localisation will

[8] OECD: https://oecd.ai/wonk/oecd-and-g7-artificial-intelligence-initiatives-side-by-side-for-responsible-ai

undercut AI innovation, and the ability for some countries / geographies to benefit from AI services offered via the cloud and unnecessarily increase costs of duplicative IT services and infrastructure.

Singapore's digital economy agreement with Australia, the Singapore-US Joint Statement on Financial Services Data Connectivity and the Joint Statement of Intent on Data Connectivity between Bangko Sentral ng Pilipinas and MAS are examples of how regulators can support free cross-border data flows.

Principle 7

Regulators should engage with the industry on areas that need further discussion

There are some areas in relation to AI that would benefit from further engagement between the regulators and the industry in order to minimise regulatory and legal uncertainties on how to apply the existing regulatory framework to the use of AI or address its risks. This engagement should be collaborative amongst regulators within a jurisdiction as well as at regional or global level to ensure alignment and consistency. Any guidance, as needed, should be proportionate to the risks posed and balance the costs and benefits so that the use of AI is not unnecessarily impeded. Conflicts with or duplication of existing regulation and requirements should be avoided.

Some areas where we would welcome further engagement include:

- **Common definition of AI:** currently there is no standard or commonly accepted definition of AI that is shared amongst regulators or international bodies or organisations. Whilst we recognise that an internationally accepted AI definition might be difficult to achieve due to local nuances, we would welcome a commonly understood term which will also facilitate the industry to understand what would be in-scope for any targeted guidance or specific new regulatory requirements that might be introduced. Any definition should be practical (e.g. avoiding the inclusion of other non-AI analytics technologies), future-proof (e.g. considering the pace of innovation in the field), broadly harmonised across major jurisdictions and compatible with the approaches of as many countries as possible. We recognise and are supportive of the work that regulators within APAC, as well as international bodies such as the Institute of Electrical and Electronics Engineers (IEEE)[9], International Organisation for Standardisation (ISO)[10] and OECD are doing to come towards a commonly understood definition and hope the GFMA's suggested definition of AI[11] can be adapted;
- **AI ethics:** We would welcome further engagement around ethical AI standards and recommend that any further guidance, if deemed necessary, is developed in partnership with the industry and in a globally consistent way to avoid any duplication.
- **Bias:** A distinction should be made between bias and unjust/unlawful bias and regulatory focus should be on unjust/unlawful bias only. Indeed, bias is not necessarily always undesirable. For example, in credit approval, an AI application discriminating based on a client's financial position can be justified but discriminating based on other personal characteristics may not be permitted or may lead to outcomes that demonstrate some kind of underlying bias. Similarly, the use of AI in marketing of products to clients or suitability assessments should avoid unfairly biased outcomes but should not prevent firms from being able to distinguish between different clients' situations and needs. We suggest that a good example is the Singapore MAS FEAT Principles definition of what it means to avoid unintended bias as "individuals and groups are not systemically disadvantaged through AI driven decisions, unless those decisions can be justified". The concept of justifiability should draw reference to general fair treatment and anti-discrimination rules that already exist. There should be recognition that any

[9] https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e_glossary.pdf
[10] https://www.iso.org/standard/74296.html
[11] Systems that act in the physical or digital world by perceiving their environment through data acquisition, interpreting the collected data, reasoning on the knowledge, or processing the information derived from this data and identifying the best action(s) to take to achieve the given goal. AI systems adapt themselves or their own algorithms by analysing how the environment is affected by previous actions, knowledge or data

guidance is not intended to be definitive given that bias definitions are not universally aligned, interpreted and enforced across regions (e.g. common law relationships between two humans recognised in one jurisdiction for calculation of income and serviceability, but not in another jurisdiction); and

- **Level of explainability and transparency for different use cases:** the level of transparency required for any AI application will be dependent firstly on the needs of the stakeholders involved and secondly on the intent or scope of the application, and will therefore be highly variable. Indeed, different use cases (e.g. loan applications versus financial crime versus marketing lead generation) warrant different levels of explainability and transparency, including between transparency to customers (e.g. through targeted disclosure) and regulators (e.g. through regulatory disclosures). We suggest that regulators support such risk-based approach. Firms should be allowed to make a risk assessment and determine the appropriate level, rather than mandating a single standard. Furthermore, rather than focusing on explainability, it is important for regulators to consider whether other methods, instead of explainability, could provide more "reasonable" or "meaningful" transparency to end-users or individuals.

3 Thematic Review of Regulatory Frameworks Supporting AI

We have described the existing general regulations and guidelines across 10 key focus areas which relate to the key risks AI poses in financial services. For each of these focus areas, we assessed how existing general regulations and guidelines pertaining to financial institutions address these risks, and whether there are any gaps.

As mentioned above, AI has the potential to create significant efficiencies and benefits for the financial services industry and its clients but at the same time might introduce or amplify certain risks. We have identified a number of key recurring focus areas in relation to AI in financial services:

- **Fairness**
- **Governance and Accountability**
- **Transparency**
- **Data Quality**
- **Data Protection**
- **Model Governance**
- **Resilience**
- **Cybersecurity**
- **Third Party Risk Managment**
- **Expertise**

In what follows, we will outline the risk backdrop for each of these focus areas, how existing financial sector regulations address the associated risks and whether there are any gaps that need to be addressed. In Section 4 we will provide an overview of AI-specific guidelines and assess these against our regulatory principles.

3.1 Fairness

Introduction of fairness in AI

Generally, fairness consists of three key aspects when using AI for decision-making: justifiability, avoiding discrimination and minimising unintended bias. Justifiability refers to the ability of firms which use AI-decisioning to produce sufficient evidence to support why a decision was made. Avoiding discrimination means the AI-application should not by design discriminate unjustifiably or illegally, for example on gender, race or ethnic backgrounds. Avoiding unintended bias refers to individuals and groups not being systemically disadvantaged through AI driven decisions, unless those decisions can be justified (MAS FEAT principles).

Risk backdrop

Fairness is integral to prevent risks that can arise when algorithms and learning models receive data that presents ingrained human flaws and biases, or when human decisions could make an algorithm discriminate unfairly. Fairness in AI protects financial institutions from deploying technology that undermines their codes of conduct and ethical values and conflicts with social values. Whilst humans can learn not to act on, or even acknowledge, such traits, AI must have fairness built into the technology and algorithms and learning models should be designed in such a way to recognise flaws and biases in input data so as to achieve fair outputs, to the extent relevant to the use case.

If not, the most notable risks that can arise are:

- **unjust bias**
  when a model unfairly skews the results of an AI model to the detriment, or advantage, of a particular group. This can occur as a result of the underlying data or as part of a model development process that gives rise to bias. For example, if facial recognition software is unable to recognise the faces of certain ethnicities not used in the development of the tools and, in the case of underwriting models, the rejection or the slowing down of the process for minority customers to acquire credit;
- **undermine an organisation's ethical values.**

Existing fairness requirements in existing general regulations and guidelines

In most jurisdictions, the regulatory standards and codes of conduct to which financial institutions are subject, include the obligation to act in the interests of customers and, to a lesser extent, the integrity of the market. Firms are also expected to follow existing laws on anti-discrimination and/or privacy laws governing the use of personal data, whether for an AI or non-AI system.

We also note that there might be interaction between fairness and other regulatory focus areas, specifically anti-discrimination laws, general "treating customers fairly" charters, sustainable finance and financial inclusion. In these instances, the aim may be to use ethical AI-algorithms to provide efficient and fair access to financial services, for example credit, to historically under-served groups.

Regulatory gaps (if any) around fairness that are specific to AI

It continues to be appropriate that laws on anti-discrimination and fairness apply broadly rather than to specific technology. Certain aspects of financial regulation that pre-date AI do impose obligations on financial intermediaries to act fairly, but this has not been extended to the very specific meaning it has been given in relation to the use of AI. There are therefore a number of areas where the industry would welcome further engagement with regulators, including:

- Concepts such as justifiability, fairness, bias, human centricity, and unequal treatment should be more clearly defined with proper consideration of these concepts and when they are compromised. Consistent use of these terms should be sought across jurisdictions. Slight variations of such harder to agree upon concepts can suggest material differences and might create a fragmented landscape making it difficult to operationalise across borders;
- Regulator-industry-academic partnerships to create guidance on testing methodologies for fairness would be useful (e.g. the Hong Kong Monetary Authority (HKMA) guidance on testing for fairness in its white paper on "Reshaping Banking with Artificial Intelligence", the MAS-led Veritas consortium); and
- Regulator-industry partnerships to develop more usable and unbiased (i.e. "fair") data sets for testing would be useful, particularly representing groups that are unique in APAC financial services.

3.2 Governance & Accountability

Introduction of governance and accountability

Governance has long been embedded in corporate and business culture encompassing rules and processes to manage, operate and control a company and its activities. AI governance requires a framework and structure governing the entire AI process. The principle of governance is intertwined with the accountability principle. Good governance requires institutions to proactively and holistically manage and monitor the use of AI, and to have appropriate mechanisms in place to ensure accountability for the development of AI, the results and impact of using AI, and the associated risks. The function is often conducted via risk management and monitoring by board and senior management, or other appropriate governance structures, e.g. specific committees or task forces, but there is also a need of responsibility for all that are involved in the AI process.

Risk backdrop

Notable risks in relation to governance and accountability that can emerge from the use of AI include:

- **Failure to monitor use of AI effectively**
  AI is a tool whose implementation needs to be monitored and overseen. Financial institutions need to ensure that the use of AI is governed by qualified officers in an appropriate control framework.

- **Regulatory breaches**
  Financial institutions which use AI but fail to put in place an appropriate governance framework are exposed to higher risks of causing harm to others who might suffer from unfair or unethical outcomes from the use of AI and regulatory breaches. If AI decisions go wrong and customers or counterparties suffer a loss or markets are impaired, the financial institutions concerned may be found negligent if they fail to have an adequate governance framework.

- **Difficulty in allocating responsibility**
  The problem of an "accountability gap" arises when it is unclear who is accountable and responsible for AI decisions. Financial institutions should implement a control framework with clear allocation of risk ownership, risk monitoring and risk assurance.

Existing governance and accountability requirements in existing general regulations and guidelines

Strong governance and oversight of the use of technology is key for ASIFMA members. In many jurisdictions, specific requirements already exist to ensure that management have full coverage of the firm's activities as well as the appropriate skillsets to perform their oversight. Specifically, financial institutions are already subject to well-established and widely understood technology risk management guidelines (such as the MAS Technology Risk Management Guidelines and the HKMA General Principles for Technology Risk Management). These guidelines already require the board of directors and senior management to ensure that a sound and robust technology risk management framework is established and maintained to manage technology risks in a systematic and consistent manner, including a clear allocation of the roles and responsibilities in managing technology risks.

At the international level, there are the Basel Committee on Banking Supervision (BCBS) Principles for the Sound Management of Operational Risk[12] and the Principles for Operational Resilience[13].

[12] BCBS (2021): Revisions to the Principles for the Sound Management of Operational Risk (bis.org)
[13] BCBS (2021): Principles for operational resilience (bis.org)

Additionally, regulators have been introducing individual accountability regimes, that aim to reduce harm to investors and strengthen market integrity by requiring firms to identify individuals that are primarily responsible and personally accountable for managing an organisation's operations and control functions, including information technology. The most senior people in financial institutions who perform key roles are typically accountable for the overall performance of the firm. This accountability extends to the actions and outcomes of AI models. Examples of such individual accountability regimes include the Hong Kong Securities and Futures Commission (SFC) Manager-in-Charge regime, the HKMA s 72B Banking Ordinance managers regime, the MAS Guidelines on Individual Accountability and Conduct, the Australian Financial Accountability Regime and the UK Senior Manager Regime. Under the UK regime, senior managers are ultimately accountable for the activities of the firm, and there is a certification requirement for staff responsible for algorithmic trading. In the US, the Federal Reserve System (FRS) places emphasis on the accountability of staff other than senior management as well. It requires that the roles in model risk management should be divided among ownership, controls, and compliance. Policies should identify the roles and assign responsibilities within the model risk management framework with clear detail on staff expertise, authority, reporting lines, and continuity.

- Regulatory gaps (if any) around governance and accountability that are specific to AI

According to the IOSCO industry engagement as part of their June 2020 consultation[14], firms implementing AI and ML mostly rely on existing governance and oversight arrangements to sign off and oversee the development and use of the AI 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 algorithms were generally not regarded as fundamentally different from non-AI algorithms and few firms identified a need to introduce new or modify existing procedural controls to manage specific AI risks. We agree with this assessment.

3.3 Transparency

Introduction to transparency

Transparency in AI refers to the level and quality of disclosure provided regarding the application of AI in services and/or products, including the risks that may be involved in AI usage. One of the key aims in promoting transparency is to instil consumer confidence and trust in AI. Transparency refers to a clear and risk-based understanding of: (i) the assumptions made in the development of AI and (ii) how AI is tested both as part of its initial development and on an ongoing basis [15].

Discussion of achieving transparency in AI often focuses on technical explainability and 'explainable' AI. Explainability typically refers to the extent to which workings of a model can be understood.

The following principles reflect regulators' expectations in relation to transparency in AI:

- Data subjects should be informed about the use of AI including when a service is powered by technology, what data is used, how it is used and the potential consequences;
- Data subjects should have access to a fair challenge and redress mechanism;
- Consumers should be educated in relation to AI; and
- Data subjects should be provided with clear explanations of what and how data was used in AI decision making.

We observe that there is a broad range of related terms

[14] IOSCO (2020): https://www.iosco.org/library/pubdocs/pdf/IOSCOPD658.pdf
[15] AFME (2019): https://www.afme.eu/Portals/0/DispatchFeaturedImages/100919%20AI%20transparency%20paper%20FINAL.pdf

for AI transparency, which are defined differently by different writers on the topic. For example, the European Commission High Level Expert Group's "Guidelines on Trustworthy AI" make reference to 'transparency', 'traceability', 'explainability' and 'interpretability'. Similarly, the "Declaration on Ethics and Data Protection in Artificial Intelligence" by the International Conference of Data Protection & Privacy Commissioners refers to 'transparency', 'intelligibility' and 'reachability'. There are no globally agreed definitions for these terms, and often one term (e.g., transparency) is used as an umbrella term for the others, encompassing a broad range of activities and principles in order to ensure that AI is transparent to consumers. In APAC, regulators have used various terms, similar to those in other regions. For example, in Singapore, MAS has used the term "transparency" and, in Hong Kong, the HKMA has referred to explainability, transparency, and provability. As mentioned above, we see transparency and explainability as being distinct.

Risk backdrop

As with any technology project, the use of AI within a firm will involve a wide range of stakeholders. Adoption of AI will be dependent on a range of stakeholders' abilities to gain and maintain trust in the firm's ethical and responsible use of the technology, even though their technical understanding of AI may vary.

This means that a lack of transparency can adversely affect:
The ability to demonstrate a suitable basis for the firm's management to sign off and oversee the firm's use of AI;
The ability to assure internal users of the application of its benefits and performance;
The ability to address concerns that might otherwise be voiced (and may indeed be voiced later) by external users or data subjects about their interaction with the firm's AI applications;
Compliance with ethical and regulatory obligations; and
Oversight, auditability and challenge by control functions, e.g. compliance, risk and internal audit.

The right level of transparency depending on the need of the stakeholders therefore helps address several risks that can arise for the financial services industry when deploying AI. We note that in reference to the risk-based approach referred to above, there needs to be a discussion of scoping operational requirements to reflect practical constraints and trade-offs that different standards of explainability and reproducibility would impose, In particular, avoid setting artificially high standards that far exceed current expectations for human-based systems.

Existing transparency requirements in existing general regulations and guidelines

Transparency is addressed by those jurisdictions that have privacy laws. For those jurisdictions that do not yet have privacy laws, this issue should be addressed with new privacy laws while reducing fragmentation and promoting interoperability; however, this issue is not specific to AI and so a specific AI regulation on transparency is not recommended.

Regulatory gaps (if any) around transparency that are specific to AI

More engagement between the public and private sector on how best to ensure there is transparency in AI use in capital markets is welcome, but care should be taken to not restrict the use of the technology as it develops. Currently, there is a specific focus on explainability, as with the HKMA and the MAS guidelines referred to above with the suggestion that AI models are either explainable or not explainable at all. We believe that such a binary approach is not appropriate for categorising AI, as it does not allow for ongoing developments in either models or explainability techniques

and it may not be able to keep pace with developments in the technology.

Instead, we suggest that a wider transparency framework is a more suitable solution. Such a framework allows suitable oversight and control of the AI model throughout its lifecycle and can be tailored to ensure that it gives the right level of detail for the different stakeholders and purposes. We suggest that such a framework should be tailored to the stakeholders in any given AI project and then built around two key elements: (i) assumptions and (ii) testing. Both should be articulated at the start of any AI project, then monitored and adjusted as necessary throughout its lifecycle. This risk-based approach can be tailored to the risk profile of each individual AI application, rather than applying 'one size fits all' standards; for example, a regulator may require more detailed information than a client. We also suggest that mandating a certain level of accuracy and validity of explainability is likely to unnecessarily limit the use of the technology, by restricting the breadth and complexity of AI models that can be used, and could also lead to the provision of 'explanations' that may be misleading and therefore counterproductive. It may also have a counterproductive effect on fairness where an AI system performs well across the general population, but may not perform well among certain subgroups. This would require further examination about the trade-offs between accuracy and fairness.

A focus on transparency instead of explainability will allow a firm to demonstrate how the AI application has been developed, how it will be used and monitored, and how it can stand up to scrutiny and challenge. Within these broad themes, transparency should meet the varied needs of individual types of stakeholder, both inside and outside the firm.

We agree with IOSCO's statement in their 2020 consultation that: "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." Indeed, transparency required for any AI application will be dependent on the needs of the stakeholders involved and will therefore be highly variable. Any regulatory guidance should allow firms to make a risk assessment and determine the appropriate level, rather than mandating a single standard.

3.4 Data quality

Introduction of data quality

In most AI applications, high data quality (i.e., data that is fit for purpose), along with high data volume, is of critical importance to ensure successful operation of AI systems. While data is an important part of training a model, data sets often naturally reflect the imperfections of the real world. It is possible to address shortcomings in training data—such as data scarcity, low quality data, and unbalanced data—through techniques like careful problem formulation, targeted sampling, synthetic data, or building constraints into models.

This entails stipulating measures to ensure that the data used in, and for training of, AI applications is fit for its intended uses in operations and decision making.

AI is reliant on balanced, high-quality data sets, and this is important for:

- The initial design of any AI application, including the establishment of any necessary parameters and rules within which it must operate; and

- The ability to "train" and test the application. This includes the ability of AI models to adapt their activity based on new data, which is critically important in the context of the AI lifecycle.

Risk backdrop

The reliance of AI on large data sets creates a dependency of the AI application on the quality of the data it is given. Where that data is inaccurate, biased or not representative of a sufficient sample size, the AI application may produce results that are unfair, inaccurate or incorrect. This is a key consideration for capital markets firms where the data they use for AI applications may pertain to clients and client activities. Poor data sets will also lead to poor AI. AI is based on a high volume of data being fed into the system and its performance ultimately depends on the quality of the data. A lack of quality data would inhibit the development of AI.

Existing data quality requirements in existing general regulations and guidelines

Most regulatory authorities make use of non-legally binding principles and guidelines to encourage institutions to ensure data quality is addressed in their businesses. Principles-based guidance on ethical guidelines are examples of how regulatory bodies address data quality issues. Both APAC and non-APAC jurisdictions adopt this approach on data quality. Notably, in the US, the FRS and Office of the Controller of the Currency (OCC) have jointly published the Supervisory Guidance on Model Risk Management, which states that the data and other information used to develop a model are of critical importance and there should be a rigorous assessment of data quality and relevance, and appropriate documentation.[16]

The BCBS in its "Principles for effective risk data aggregation and risk reporting"[17] lay out that the board and senior management should promote the identification, assessment and management of data quality risks as part of its overall risk management framework and that supervisors expect banks to measure and monitor the accuracy of data and to develop appropriate escalation channels and action plans to be in place to rectify poor data quality.

In Australia, the Privacy Principle Guidelines[18] number 10 states that firms must take reasonable steps to ensure that the personal information it collects is accurate, up-to-data and complete and to ensure that the personal information it uses and discloses is – having regard to the purpose of the use or disclosure – accurate, up to date, complete and relevant. Similarly, in Hong Kong, the Office of the Privacy Commissioner for Personal Data's (PCPD) principle number 2 requires that data users should "take all practicable steps to ensure that personal data is accurate". [19] In Singapore, the Personal Data Protection Act[20] has an accuracy obligation which requires an organisation to make a reasonable effort to ensure that personal data collected by or on behalf of the organisation is accurate and complete, if the personal data is likely to be used by the organisation to make a decision that affects the individual to whom the personal data relates; or is likely to be disclosed by the organisation to another organisation. In Korea, the Personal Information Protection Act[21] article 3 requires that the personal information controller shall ensure personal information is accurate, complete, and up to date to the extent necessary in relation to the purposes for which the personal information is processed.

[16] FRS (2011): https://www.federalreserve.gov/supervisionreg/srletters/sr1107a1.pdf.
[17] BCBS (2013): Principles for effective risk data aggregation and risk reporting (bis.org)
[18] Office of the Australian Information Commissioner (2019): Australian Privacy Principles guidelines – OAIC
[19] PCPD (1996): https://www.pcpd.org.hk/english/data_privacy_law/ordinance_at_a_Glance/ordinance.html
[20] PDPC (2017): https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Advisory-Guidelines/the-accuracy-obligation---ch-16-(270717).pdf
[21] Korean Legislation Research Institute (2014): https://elaw.klri.re.kr/eng_service/lawView.do?hseq=53044&lang=ENG

Regulatory gaps (if any) around data quality that are specific to AI

We suggest that when jurisdictions consider issuing data quality governance guidance, this should not be limited or specific to AI. Any such guidance should also follow international standards, such as BCBS Principle 239 on effective risk data aggregation and risk reporting.

Similarly, data privacy regulations, that are often in the process of being updated by non-financial services regulators, should not aim to be limited to a particular industry or technology use case and should remain principles driven.

3.5 Data Protection

Introduction to data protection

The use of personal data and sensitive personal data[22] and a financial institution's ability to abide by data protection laws when using such personal data in AI models should be assessed. Some of the more recent data protection laws impose limitations on the use of AI and express data subjects' rights with respect to the use of their personal data in AI models and the impacts AI-driven decisions may have on them. AI also introduces data protection risks where, as a developing technology, its use may undermine controls being used to mitigate data protection risk. For example, institutions should be attentive to the possibility of running afoul of data privacy laws by re-identifying personal data as a result of the insights AI models can draw from using different de-identified data sets.

The following principles reflect regulators' expectations in relation to data protection and the use of AI:

- Ensure compliance with applicable personal data laws and pay regard to relevant good practices. There should be lawful usage and protection of personal data while providing AI-enabled services.
- Implement policies and controls to protect personal data in connection with AI specifically. The policies and controls should consider mechanisms, such as:
  - fair and lawful data collection;
  - data minimisation;
  - disclosure and/or informed consent;
  - controlling access to networks, web applications and client applications;
  - classifying information into various sensitivity levels; and
  - implementing rules for collecting, storing, processing and using individuals' information.
- Adopt appropriate anonymisation and de-identification methods to help protect data privacy.
- Ensure that the rights of individuals with respect to how they can control their personal data are adequately protected in relation to the use of AI.

Risk backdrop

The principle of data protection in AI can guard against the risks of unauthorised collection, sharing or use of personal data that can arise when using AI to collect or use such data. For example:

- insufficient authority to use/breach of consent – financial institutions collect personal data from customers, including biometric data, investment preferences, backgrounds and often family information. This is often provided under specific terms (and sometimes consent) for a specified purpose. If the data collection and use mechanisms are not properly designed and (where relevant) disclosed so that customers can provide proper consent where necessary, then customers' privacy rights may be at risk. Consent should be specific and clear. Secondly, if adequate controls are not deployed, institutions may exceed the scope of the relevant

[22] The definitions of personal data and sensitive personal data vary from jurisdiction to jurisdiction https://www.asifma.org/wp-content/uploads/2020/07/asifma-jurisdictional-comparison-grid-of-data-protection-rules-v20200721-final.pdf

disclosure and/or consent. Using personal data for purposes other than the purpose for which consent has been given, can lead to litigation and regulatory action as well as loss of reputation; and

- data privacy breaches – while not specific to AI, we note that if a greater volume and variety of data is held centrally as a result of the implementation of AI, then this increases both the risk (i.e., honeypot) and consequences of any breach. A security breach can lead to the theft of personal data (e.g., biometrics) that can never be replaced.

Existing data protection requirements in existing general regulations and guidelines

The underlying data protection questions for even the most complex AI project are generally much the same as with any new project. (e.g., Is data being used fairly, lawfully and transparently? Do people understand how their data is being used? Is data being kept secure? Do people retain adequate control over their data?)

Most advanced economies have robust data privacy regulations, with a number of markets (including a number of emerging economies in APAC) bolstering their standards following the EU's issuance of the GDPR. This includes the 2020 update to the Korea Personal Information Protection Act[23]. In July 2020, ASIFMA published a jurisdictional comparison grid of data protection rules[24]. Most of these data protection laws reflect long-established principles. The 1980 OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data articulate eight basic principles of data protection: collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation and accountability. In Japan, the Personal Information Protection Commission (PPC) Bill to Amend the Act on the Protection of Personal Information introduces the concept of pseudonymised information and require the publication of its purposes of use[25]. In the US, sections 501 and 505(b), 15 U.S.C. 6801 and 6805(b), of the Gramm- Leach-Bliley Act[26], as implemented through interagency guidelines establishing information security standards, address standards for developing and implementing administrative, technical, and physical safeguards to protect the security, confidentiality, and integrity of customer information.

However, more precise standards are evolving to recognise the choice individuals should have when data concerning the individual is used in AI– for example, Article 22 of the GDPR contains specific protections relating to decisions based solely on automated processing, including profiling, where it produces legal effects concerning a person or similarly significantly affects them.

Regulatory gaps (if any) around data protection that are specific to AI

As financial institutions increasingly leverage and collect various data sources that will be used for AI purposes, they should create robust data protection frameworks. It is recommended that policymakers/regulators identify and focus on issues relating to data privacy/protection when using AI, for example around anonymisation and pseudonymisation: anonymised data is generally not subject to data privacy/protection laws.

[23] Personal Information Protection Commission (2020): http://www.pipc.go.kr/cmt/english/news/selectBoardArticle.do
[24] ASIFMA (2020): https://www.asifma.org/wp-content/uploads/2020/07/asifma-jurisdictional-comparison-grid-of-data-protection-rules-v20200721-final.pdf
[25] IAPP (2020): https://iapp.org/news/a/japan-enacts-the-act-on-the-protection-of-personal-information/
[26] US Government (1999): https://www.govinfo.gov/content/pkg/PLAW-106publ102/pdf/PLAW-106publ102.pdf

3.6 Model Governance

Introduction of AI model governance

AI model governance refers to a broader scope of oversight that covers the full model development cycle from proposal to deployment to ongoing outcome analysis and governance.

Regulators generally expect that:

- There should be appropriate processes for model development, testing and independent validation to ensure that the model fulfils the model design objectives with an appropriate level of human involvement throughout the process of developing and maintaining AI models;
- Risk management and controls should be used throughout the model development lifecycle process to ensure that AI algorithms comply with all relevant regulatory principles and documentation of the model is kept to enable consistent algorithm model and transparency; and
- Models be regularly refreshed using updated training datasets that incorporate new input data as commercial objectives, risks or corporate value changes.

Risk backdrop

The lifecycle of an AI model is lengthy, from proposal to development, implementation and ongoing governance. The deployment of AI is an ongoing, continuous process, rather than a one-off implementation. It is critical for financial institutions to monitor the deployment of AI applications.

Risk relevant to model management include:

- Decline in AI performance: It is possible that the performance of AI models may decline as, more than a typical model, they rely on not only historic data but also data they are processing, and they may be sensitive to certain variables or assumptions that are liable to change which may lead to the assumptions and results of AI solutions becoming inaccurate over time;
- Unintended outcomes: AI models rely on their data inputs, which can incorporate bias, may have gaps or data quality problems and do not meet real life expectations (i.e. unrepresentative data as opposed to discrimination). This may be a particular risk when there is inadequate human supervision of the outcomes. While checking bias in the input data is a good first step to avoid unintended outcomes, it is also important to check if there are biases introduced during the model training, either due to incorrect optimisation parameters or algorithm;

Existing model governance requirements in existing general regulations and guidelines

In Hong Kong, the HKMA[27] and the SFC provided thematic guidance on algorithmic trading and robo-advisory algorithms covering requirements for algorithm supervision and testing, as well as risk management and controls. Convergence of control frameworks for AI governance and algorithmic trading must be considered as institutions start to utilise AI in their trading algorithms. The HKMA, requires regular internal and external audits for algo-trading activities only, and this is not specific to AI algorithms.

Korea focuses on algorithmic trading and requires that the trading systems must be validated, tested and approved prior to launch and immediate measures must be taken when trading goes beyond established parameters[28].

Australia provides guidance applicable to offering "digital advice"[29] (i.e. robo-advice) and emphasises keeping appropriate documentation during system design, testing the strategy, change management processes, security,

[27] HKMA (2020): https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2020/20200306e1a1.pdf
[28] Available at http://www.smallake.kr/wp-content/uploads/2016/01/20160124_111425.pdf (in Korean)
[29] Digital advice refers to automated financial product advice using algorithms and technology and without the direct involvement of a human adviser

and control over algorithm change as well as the ability to suspend advice if an error is identified. Such regulations are not unique to AI models, and would be applicable to even non-AI models.

Supervisory guidance from the US FRS and OCC[30] emphasise that depending on materiality, ongoing monitoring may be required to evaluate whether changes in products, exposures, activities, clients, or market conditions necessitate adjustment, redevelopment, or replacement of the model, which we take to include AI models given the definition of model[31] and to verify that any extension of the model beyond its original scope is valid. Any model limitations identified in the development stage should be regularly assessed over time, as part of ongoing monitoring. Monitoring begins when a model is first implemented in production systems for actual business use. This monitoring should continue periodically over time, with a frequency appropriate to the nature of the model, the availability of new data or modelling approaches, and the magnitude of the risk involved. Banks should design a program of ongoing testing and evaluation of model performance along with procedures for responding to any problems that appear. This program should include process verification and benchmarking.

Regulatory gaps (if any) around model governance that are specific to AI

As mentioned above, traditionally, financial services regulations have governed the use of algorithms that are models reasonably well. Because AI systems pose risks that are similar to those posed by quantitative models generally, most financial institutions leverage existing model risk management frameworks. However, the application of certain AI models may present certain risks that non-AI models do not exhibit, for example the ability to change behaviour over time and the reliance on data more than human inputs or reasoning. This is in relation to their ability to drift off their original model because they are learning through the application of AI so they may evolve beyond their original rule set that they are designed for. This requires monitoring to ensure this does not happen or happens inappropriately. This is a type of monitoring that normal algorithmic model governance does not regulate explicitly for and financial institutions should consider updating their model governance policies to address any unique or heightened risks posed by AI models.

3.7 Resilience

Resilience aims at helping to improve the stability and reliability of services relating to AI in order to build trust and confidence in financial institutions for their customers, counterparties and/or markets. Issues such as disruption in data quality, unintended failures in development and erroneous recommendations can destroy trust and expose financial institutions to risk.

Many jurisdictions including Hong Kong, South Korea, Thailand and the EU have put resilience as one of the core principles in guidelines related to AI.

Building and maintaining resilience in the use of AI is a top priority for financial institution users. Like any other connected system, a connected AI system should be subject to existing and developing resilience guidelines.

Existing resilience requirements in existing general regulations and guidelines

Financial institutions are already subject to business continuity management guidelines and there is an increasing regulatory focus globally on operational resilience including the 2021 UK Prudential Regulation Authority

[30] FRS (2011): https://www.federalreserve.gov/supervisionreg/srletters/sr1107a1.pdf
[31] Model: a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques and assumption to process data into quantitative estimates

(PRA), FCA and BOE policy on "operational resilience: impact tolerances for important business services" [32], the 2021 BCBS principles on operational resilience and the 2020 US FRS, OCC and Federal Deposit Insurance Corporation "Sound practices paper on operational resilience" [33]. In Hong Kong, the HKMA is considering the need to provide additional guidance to implement the BCBS Operational Resilience principles [34].

Additionally, some regulators have additional guidance in relation to resiliency and business continuity around algo trading systems. For example, in Hong Kong, the Code of Conduct issued by the SFC [35] requires the following measures in relation to electronic trading services for algorithmic trading systems:

a. a written contingency plan to cope with emergencies and disruptions related to such services and the performance of regular testing of such contingency plan. The objectives of such requirements are to ensure that the licensed corporation can rectify any such emergency situation in a timely manner and inform clients and stakeholders the possible causes and how orders are being handled; and

b. for a licensed corporation to implement effective controls to immediately prevent the electronic trading system from generating and sending orders to the market and cancel any unexecuted orders. Such resilience is referred to as "kill-switch" functionality that can shut down algorithm in the event of an emergency.

As disclosed in the Appendix to the "Circular to all Licensed Corporations on Algorithmic Trading"[36] issued by the SFC on 13 December 2016, licensed corporations have largely implemented emergency trade flow shutdown controls at sufficiently granular levels (e.g. at the exchange connectivity, algorithmic, order, trader, system and client levels) to control the level at which trade flows can be shut down in a timely manner in an emergency. The SFC further notes that, whilst licensed corporations may have contingency plans for continuing and recovering critical business functions under various crisis situations, they are required to formulate written contingency plans including details of procedures to cope with different emergency scenarios such as shutdown controls at different levels to provide specific responses with appropriate timelines.

Regulatory gaps (if any) around resilience that are specific to AI

Most APAC jurisdictions regulate for business continuity but not yet for operational resilience, but this is expected to change following the 2021 finalisation of the BCBS Principles on Operational Resilience. Operational resilience focuses on the continued ability to provide critical services and products to customers, counterparties and markets from an end-to-end operational process perspective even when those processes extend to and rely on outside third parties. As mentioned above, global standard setters such as the BCBS are increasingly focused on operational resilience and it is expected that APAC jurisdictions will follow. The operational resilience of AI will be a matter that will have to be considered when operational resilience rules are developed.

[32] PRA, FCA, BOE (2021): PS6/21 | CP29/19 | DP1/18 Operational Resilience: Impact tolerances for important business services | Bank of England
[33] OCC (2020): Agencies Release Paper on Operational Resilience | OCC
[34] HKMA (2021): Principles for Operational Resilience and Revised Principles for Sound Management of Operational Risk (hkma.gov.hk)
[35] SFC (2020): Microsoft Word - Code_of_conduct Sep 2020_Eng 2nd.doc (sfc.hk)
[36] SFC (2016): https://www.sfc.hk/edistributionWeb/gateway/EN/circular/intermediaries/supervision/doc?refNo=16EC67

3.8 Cybersecurity

Introduction to cybersecurity in AI

Using HKMA regulatory guidance as an example, cybersecurity refers to the ability to protect or defend against cyber-attacks [37]. Cybersecurity risk is a growing global challenge, and one that governments are addressing with several initiatives. These include issuing regulations, guidance and supervisory practices, in order to set up security frameworks that mitigate security risk, which included ensuring security in the use of AI.

Risk backdrop

We suggest there are nuanced but important differences between cybersecurity for AI and cybersecurity for I.T. systems and networks. In summary, cyberthreats for I.T. systems and networks primarily aim to steal resources e.g., hold data hostage, exfiltrate/steal confidential data, deny uses of computer systems, send fake instructions. Cyberthreats for AI aim to manipulate the outcomes of AI models e.g., poison data, influence data models and parameters, influence the drivers/motives of an automated action, influence and change results that are used for decisions which could cause financial harm. For example, adversarial machine learning is a class of threats that are unique to AI, and includes poisoning of training data to influence model accuracy/performance; extraction or inference of sensitive information from trained models; and a specially crafted input to cause the model to make mistakes. Mitigation against adversarial machine learning includes training data integrity protection, model robustness test, and adversarial example filtering. Another key cyberthreat (AI) is potential manipulation of an adaptive system to force it out of the ability to self-correct.

The cyberthreat environment is evolving and bad actors do leverage AI technologies to launch more sophisticated attacks as attackers can maliciously use AI to help speed-up, scale-up and target their attacks through automation and deep learning analytics to predict the victim's move. At the same time, defenders can also use AI to improve their analytical ability to predict the threat actors' next moves and to enhance cybersecurity capabilities and cyber-defence.

Existing cybersecurity requirements in existing general regulations and guidelines

Existing cybersecurity regulations and guidelines on IT systems and networks will also apply to AI development. These cybersecurity considerations should be included throughout the AI lifecycle.

Across APAC, the jurisdictions reviewed generally have in place generic regulations and guidance addressing or setting out requirements on cybersecurity.

For example, the People's Bank of China Initiative on the Effective Protection of Personal Financial Information details that financial institutions must establish cybersecurity policies including procedure, inspection, evaluation and issue handling. The initiative also stipulates that financial institutions must implement effective data protection measures and ensure cybersecurity controls, so that they are able to effectively deal with data corruption and attacks. Financial institutions must remain vigilant to emerging cybersecurity threats. The initiative also says information should be classified into various sensitivity levels, to ensure required protection is given to the most sensitive data.

The Reserve Bank of India's Pre-Paid Payments Instruments Guidelines stipulate the need for data security infrastructure

[37] HKMA (2015): https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2015/20150915e1.pdf

and systems for prevention and detection of fraudulent activity. These guidelines also specify the need for review and monitoring, information security policy as well as security incident protocols[38].

Testing banks' vulnerability and resilience to cyber-risk is another common requirement across jurisdictions. For example, the non-binding Australian Securities and Investments Commission's (ASIC) Regulatory Guide 255 asserts the need for a documented test strategy, change management processes, security, and control over algorithm changes[39].

Regulatory gaps (if any) around cybersecurity that are specific to AI

As existing regulatory frameworks for cybersecurity tend to be principles-based, reflecting the regularly changing landscape of potential cyberattacks, we do not observe that new regulations are required that are specific to AI in APAC as existing cybersecurity regulations also address cybersecurity risks in relation to AI.

3.9 Third-party Risk Management (TPRM)

Introduction to TPRM

Requirements on TPRM aim to govern the selection, monitoring and overseeing of the contractual relations of financial institutions with any third parties who provide services or functions in relation to any AI system that is used in business activities (such as developers and vendors of AI systems). Third party risk management and the management of interdependencies is closely linked to the topic of (operational) resilience.

Risk backdrop

The careful selection, monitoring and governance of third parties that provide AI-related services, functions or products, in addition to robust contractual rights can help govern the risks that arise from the engagement of a third-party for AI. Ultimately, it is the financial institution that is responsible for the effective management of the risks related to the use of third-party AI. The most significant risk arises from losses incurred if the third-party service provider fails to perform or suffers a hack or breach, for example:

- Loss-making AI powered trading/investment – certain AI tools have been developed such that investment decisions are made without any human interaction. Investment managers can contract with third parties boasting such technology to manage funds on the concept that funds traded without bias or emotion will perform better. However, if the AI is flawed, without close monitoring to quickly discover and stem losses, significant losses can be caused by inappropriate trading or investment strategies;
- data breach and misuse – certain financial institutions use biometrics for customer authentication and the technology is often provided by a third party. Private and confidential data breach and misuse (in terms of selection and handling) can occur at the third-party level. Due to the amount of sensitive data, significant losses can occur for the financial institutions. Without clear liability models, recovering such losses can take significant time and incur significant legal fees. Reputational damage for the financial institution cannot be undone; and
- risk of interdependencies and interactions between third party AI and other technology systems and AI – it is key for financial institutions to understand how their own and third party AI applications or other IT systems interact and the dependencies between them. In the absence of this, it might be difficult to determine causation and accountability for regulatory breaches and incidents.

[38] RBI: https://rbidocs.rbi.org.in/rdocs/content/pdfs/PPICC280314_A.pdf
[39] ASIC (2016): https://asic.gov.au/regulatory-resources/find-a-document/regulatory-guides/rg-255-providing-digital-financial-product-advice-to-retail-clients/

Existing TPRM requirements in existing general regulations and guidelines

Financial institutions are already subject to well-established outsourcing and third-party risk management rules and guidelines which we suggest adequately address the third-party risk in the context of AI.

For example, the HKMA[40] and the MAS[41] [42] have well established requirements on outsourcing, covering accountability, risk assessment; ability of service providers, outsourcing agreement, customer data confidentiality, control over outsourced activities, contingency planning, access to outsourced data and concerns in relation to overseas outsourcing.

In South Korea, for TPRM of core functions of financial institutions, the Financial Services Commission has permitted approved third parties to provide financial institutions core functions and to enable sandbox-approved transformative financial services.

In Australia, the Australian Prudential Regulation Authority (APRA) Prudential Standard 231 on outsourcing requires that all outsourcing arrangements involving material business activities entered into by an APRA-regulated institution be subject to appropriate due diligence, approval and ongoing monitoring. The standard 234 on information security aims to minimise the likelihood and impact of information security incidents on the confidentiality, integrity, or availability of information assets, including information assets managed by related parties or third parties.

Most jurisdictions have certain requirements on third parties including the European Central Bank in the EU, and the FRS and Securities and Exchange Commission in the US.

Regulatory gaps (if any) around TPRM that are specific to AI

Overall, the existing requirements on third party management are relatively comprehensive.
Global standard setters including IOSCO and the Financial Stability Board (FSB) are also reviewing and assessing the updating of existing outsourcing and third-party risk management guidelines considering recent technological developments and lessons learned from the Covid-19 pandemic.

3.10 Expertise

Introduction to expertise

AI performance largely depends on the talent of highly skilled individuals as well as necessary skills by the everyday user/employee at the financial institution. The ability to identify, hire, develop, motivate, and retain highly skilled personnel is one of the key success factors in an AI strategy. Competition in industry for qualified expertise is intense. Regulations and guidance in different jurisdictions also require organisations to have staff with the requisite competence and organisational controls to supervise relevant staff.

Risk backdrop

Financial institutions interested in deploying AI applications need to have competent professionals who are capable of assessing risks of AI applications and maintaining and monitoring the AI applications, otherwise risks include:

- **Poor maintenance of AI application:** Successful maintenance and application of AI require qualified and competent staff members, who are familiar with the technology used. This may be a challenge for some financial institutions. Financial institutions who do

[40] HKMA: https://www.hkma.gov.hk/media/eng/doc/key-functions/banking-stability/supervisory-policy-manual/SA-2.pdf
[41] MAS (2018): https://www.mas.gov.sg/regulation/guidelines/guidelines-on-outsourcing
[42] MAS (2021) Consultation Paper on Notices to Banks and Merchant Banks on Management of Outsourced Relevant Services (mas.gov.sg)

not have the necessary qualified staff are unlikely competent to handle and monitor their AI applications. Financial institutions need to ensure their ability to maintain their AI applications in case of staff turnover of their AI specialists; and

- **Over-reliance on third-party service providers:** Given the unprecedented pace of development of AI, it is also challenging for financial institutions to train their staff and ensure their knowledge is up-to-date. Financial institutions without sufficient in-house capability may heavily rely on external service providers but may be unable to supervise and monitor their performance and uncover issues after the fact that may be unable to be remediated.
- **Reporting and remediating errors during routine use of the application.**

Existing expertise requirements in existing general regulations and guidelines

In the current regulations and rules, there are general requirements on the resources involved in the design, development and the approval of the use of AI applications. These requirements are currently principle-based and provide room for financial institutions to determine the specific requirements of competent resources.

Regulatory gaps (if any) around expertise that are specific to AI

The regulations mainly cover the expertise related to the design and development of AI applications. The expertise required for other relevant functions such as the competency and business knowledge of the testing team (especially for teams testing AI applications developed by a third party), and compliance personnel for monitoring and ensuring compliance of the AI applications may not be covered.

We note that it is also crucially important that regulatory bodies develop the skills and resources to respond to and support the development of AI within their industries. This will also allow development of AI as a regulatory tool, for example for assessing large quantities of data or predicting the build-up of risk.

4 Jurisdictional Overview of AI-Specific Guidance

This section provides an overview of AI-specific guidelines issued to date in APAC and assesses these AI-specific guidelines against the focus areas identified in this paper. For completeness and to demonstrate global developments, we have also included a snapshot of AI-guidelines for the EU, US and UK.

4.1 Hong Kong

In Hong Kong, the HKMA in 2019 published a circular on high-level principles of AI[43] as well as "Guiding Principles: "Consumer Protection in respect of Use of Big Data Analytics and Artificial Intelligence by Authorized Institutions"[44] (Guiding Principles). The SFC as yet has not issued any AI-specific regulations. The PCPD in 2018 issued its ethical accountability framework[45].

Fairness

The HKMA AI principles include fairness as a principle. The principles require banks to ensure AI-driven decisions do not discriminate or unintentionally show bias against any group of consumers. The HKMA has further elaborated in its White Paper on Reshaping Banking with Artificial Intelligence[46], similar requirements to those set out by the MAS (see below) on fairness. Banks need to formalise an enterprise-wide governance and quality control policy and processes to evaluate data by testing the representativeness of data to ensure that an AI model achieves fair outcomes for customers. Although the HKMA did not provide any definition of fairness, it has recommended some detection tools that can assist a bank to check model discrimination. These include: (i) bias detection tools – comparing performance metrics of models across different data groups; (ii) fairness detection tools – measuring fairness by comparing the outputs of disadvantaged versus advantaged groups using a fairness definition, e.g. equal opportunity; and (iii) bias invention tools – computing the decision boundaries of fairness for discriminated groups by re-adjusting thresholds until disparities are minimised. The HKMA in the White Paper has also recommended banks to establish an AI Center of Excellence (CoE) in the whitepaper on "Reshaping Banking with Artificial Intelligence". The proposed CoE acts as a centralised function to align AI technology with broader ethical, governance and privacy regulations.

The PCPD emphasised that personal data must be collected by means which are fair and lawful, and not excessive having regard to the purposes. The models and algorithms used should mitigate bias, illegal discrimination and other inappropriate actions.

Governance and Accountability

The HKMA in its AI principles has expressed stipulations on who in an organisation should be accountable, providing that the board and senior management of institutions should remain accountable for all the decisions and processes driven by big data analytics and AI. The HKMA also highlights four key areas of good governance that the banking sector should adhere to, including having a documented governance framework, explainability, adherence to the consumer protection principles and validation.

The Hong Kong the PCPD also suggested a Process Oversight Model to promote accountability in the use of AI as part

[43] HKMA(2019): https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2019/20191101e1.pdf
[44] HKMA (2019): Consumer Protection in respect of Use of Big Data Analytics and Artificial Intelligence by Authorized Institutions (hkma.gov.hk)
[45] PCPD (2018): https://www.pcpd.org.hk/misc/files/Ethical_Accountability_Framework.pdf
[46] HKMA (2019): https://www.hkma.gov.hk/media/eng/doc/key-functions/finanical-infrastructure/Whitepaper_on_AI.pdf

of its guidance. The guidance covers matters including: (i) accountability for the oversight process, (ii) translation of organisational values into principles and policies, (iii) translation of organisational values into an "ethics by design" program, (iv) review according to an internal process, and (v) accountability to the individual. In this model the PCPD specifically states the principle of accountability to the individual, in addition to the institution as a whole.

Transparency

The HKMA has set out in its 2019 Guiding Principles that it expects all banks to observe transparency and disclosure principles. The HKMA expects banks to:

- increase consumer confidence by making clear when a service is powered by AI technology and the risks involved;
- provide accessible and fair complaint handling and redress mechanisms for BDAI-based products and services; and
- provide consumer education and data use disclosures.

The PCPD also requires proactive disclosure regarding AI use and consequences and expects organisations to have policies that communicate the data stewardship values that govern the AI-driven data processing activities.

Data Quality

The HKMA requires banking institutions to (i) use their data governance frameworks to ensure good quality data and relevance, (ii) assess accuracy, completeness, timeliness and consistency of data, and (iii) implement rigorous validation and testing to confirm accuracy before system deployment.[47]

Data Protection

In its 2019 AI principles, the HKMA requested banks to implement effective data protection measures, comply with the applicable personal data laws, pay regard to relevant good practices issued by the PCPD, and ensure privacy by design, data minimisation and informed consent. They also said that security controls should be able to effectively deal with data poisoning and attacks and that financial institutions should remain vigilant to emerging security threats.

Model Governance

The HKMA has provided high-level guidance on conducting periodic reviews and on-going monitoring. They state that, since AI applications can learn from live data and their model behaviour may hence change after deployment, financial institutions should conduct periodic reviews (e.g. re-validation of the AI model where appropriate) and on-going monitoring to ensure that the applications continue to perform as intended.[48] The HKMA articulated that, due to the nature of AI, new AI models are frequently updated and deployed. Financial institutions need a robust handover process between the development and operation teams to ensure the deployment process does not disrupt existing processes. This is typically done by conducting automated continuous testing before deployment.[49]

Resilience

The HKMA in its AI principles highlighted the need for risk mitigation and contingency plans in on-going monitoring and maintenance stage(s), and even the most robust AI applications may deliver unintended outcomes. Apart from subjecting their AI-driven activities to appropriate risk-mitigating controls (e.g. human-in-the-loop mechanism, prudent risk limits and sample quality assurance checks), banks should implement contingency measures that can

[47] HKMA (2019): https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2019/20191101e1.pdf
[48] HKMA (2019): https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2019/20191101e1.pdf
[49] HKMA (2019): https://www.hkma.gov.hk/media/eng/doc/key-functions/finanical-infrastructure/Whitepaper_on_AI.pdf

promptly suspend AI applications and trigger fall back procedures (e.g. human intervention or conventional processes) where necessary.

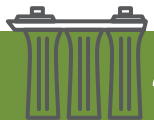
Cybersecurity

The HKMA said that security controls should be able to effectively deal with data poisoning and attacks and that financial institutions should remain vigilant to emerging security threats.

Expertise

The HKMA in its AI principles highlighted the need to have sufficient expertise in the application design and development stage:

- Given that designing and developing AI applications requires specific expertise, banks should ensure that their developers have the requisite competence and experience; and
- Senior management should satisfy themselves that there is an effective mechanism to supervise the relevant staff. They should also implement appropriate programmes to recruit, train and retain employees with suitable skillsets.



4.2 Singapore

In November 2018, the MAS published the "Principles to Promote Fairness, Ethics, Accountability and Transparency in the Use of AI and Data Analytics in Singapore's Financial Sector" (the FEAT Principles)[50]. The FEAT Principles set out the 14 principles in four categories: fairness, ethics, accountability and transparency, for financial institutions to consider when assessing or developing a framework to govern the use of AIDA.

In January 2019, the Singapore Personal Data Protection Commission (PDPC) published its "Proposed Model AI Governance Framework"[51] (the PDPC AI Framework), a sector-agnostic, crosscutting set of principles and guidelines. Both sets of guidance were developed based on consultation with industry. Notably, both the MAS and PDPC have encouraged feedback from industry and other stakeholders on implementation of the guidance documents, with a view to seeing them as "living documents" that could be amended in future.

Fairness

In the FEAT principles, the MAS expresses the concern that many of the datasets being used to train AI systems would not represent a wide population, which might lead to unfair decisions due to the use of unrepresentative datasets. In the paper, the MAS has also provided illustrative examples which are useful in guiding AIDA firms to comply with the principles of fairness. MAS' FEAT Principles also contain two key ethical standards with illustrations: (i) AI-driven decisions should be aligned with the firm's ethical standards; and (ii) AI decisions are held to at least the same ethical standards as human driven decisions. MAS also launched the Veritas Initiative which is a good example of public-private partnership to develop and test fairness metrics and tools against financial services use-cases.

The PDPC also encourages a human-centric approach for AI governance with the focus being on protecting the well-being and safety of human beings in the design, development and deployment of AI. While the human-centric approach is not legally binding on organisations, it is a tool to assist organisations to implement governance frameworks for AI.

Some of the aims of this approach include:

- equal distribution of benefits and creating possible benefit from the use of data and advanced modelling techniques;
- encouraging the practice of virtues;
- making decisions that do not cause foreseeable harm to an individual; and
- allowing users to maintain control over the data being used.

Governance and Accountability

MAS has elaborated and divided the principle of accountability into internal and external aspects[52]. For internal accountability, the MAS focuses on the approval and monitoring within organisations. The MAS requires that:

- the use of AIDA-driven decision-making is approved by an appropriate internal authority;
- firms using AIDA are accountable for both internally developed and externally sourced AIDA models; and
- firms using AIDA proactively raise management and board awareness of their use of AIDA on a high-level basis or on specific issues which arise.

[50] MAS (2018): https://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf
[51] Available at https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf
[52] MAS (2018): https://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf

The external aspect of accountability primarily involves data subjects. As general illustrations, the Singapore MAS requires that:

- data subjects are provided with channels to enquire about, submit appeals for and request reviews of AIDA-driven decisions that affect them; and
- verified and relevant supplementary data provided by data subjects are taken into account when performing a review of AIDA-driven decisions.

The PDPC AI Framework highlights internal governance structures and internal control measures as one of its four key focus areas.

Transparency

The MAS in its FEAT principles stated that transparency and explainability are some of the core principles that financial institutions should observe when offering AI-services. The MAS expects:

- proactive disclosure of AI use to data subjects as part of general communications so as to increase public confidence;
- upon request, data subjects to be provided with clear explanations of what and how data was used as part of an AIDA-driven decision about the data subject; and
- upon request, data subjects to be provided with a clear explanation of consequences that AIDA-driven processes may have on them.

The Singapore PDPC also requires proactive disclosure regarding AI use and consequences and one of its two "guiding principles" is that AI-driven decision-making processes must be transparent and explainable.

Data Quality

The PDPC AI Framework encourages organisations to ensure data quality and understand and address factors that may affect the quality of data, including accuracy, completeness, veracity, timing, relevance, integrity, usability of the dataset used and human interventions.

4.3 China

Fairness

In China, fairness was part of the "Beijing AI principles[53]" jointly developed by the Ministry of Science & Technology (MOST), the Beijing Academy of Artificial Intelligence (BAAI), together with various leading Chinese academics, and an AI industrial league involving firms including Baidu, Alibaba and Tencent. The publication set out 15 principles to be considered throughout different stages of AI application. It addressed fairness in terms of "humanity" and emphasised that the development and research of AI needs to be diverse and inclusive and consequently be fair to all parties. The MOST has stated that, first and foremost, the overall goal of AI development should be to promote the well-being of mankind and that AI should promote green development to meet the requirements of environmental friendliness and resource conservation.

Governance and Accountability

The MOST has expressly included the principle of accountability in its "Governance Principles for the New Generation Artificial Intelligence" [54] (Governance Principles). There should be accountability mechanisms for AI and responsibilities of AI developers, users and the impacted parties should be clarified.

Transparency

China's MOST has adopted two unique positions within the region:

- the transparency principle has been expanded to encourage the establishment of an AI open platform to avoid data/platform monopolies; and
- transparency should be "continuously" improved.

Data Quality

Whilst data quality is not expressly and individually stated in its local guidance, it is inherently enshrined in the principles of fairness and justice (to eliminate prejudice in data acquisition) and inclusive and sharing (to avoid data monopolies) among the MOST Governance Principles.

Data Protection

China's financial regulators and privacy regulators issued various extensive initiatives on data privacy protection in relation to AI use that applies to financial institutions, stakeholders concerned with AI development and other people that are involved in collecting and using personal information. The initiatives extensively set out protection mechanisms. In addition to protection mechanisms, the MOST has also set out that there should be full protection of individuals' right to be informed and to choose.[55]

[53] BAAI (2019): https://www.baai.ac.cn/news/beijing-ai-principles-en.html
[54] MOST (2019): http://www.most.gov.cn/kjbgz/201906/t20190617_147107.htm (in Chinese)
[55] China Daily (2019): http://www.chinadaily.com.cn/a/201906/17/WS5d07486ba3103dbf14328ab7.html

4.4 India

In India, the National Institution for Transforming India (Niti) Aayog developed a "National Strategy for Artificial Intelligence" [56].

Fairness

The Niti Aayog strategy included fairness as one of the key topics in the use of AI applications and aims to tackle potential biases.

Governance and Accountability

Security and Exchange Board of India has issued a number of legally binding circulars which require financial institutions, to submit regular reports about the offering and the use of AI and machine learning applications and systems. In these reports, the financial institutions are required to set out details of the controls, safeguards and audit requirements in place for the AI or machine learning systems and applications.

4.5 Korea

Fairness

In Korea, the Korea Communications Commission and Korea Information Society and Development Institute jointly announced principles to govern the creation and use of AI, focusing on the protection of human dignity. In particular, the principles compel all members of society, including government, organisations and users, to acknowledge AI may cause a social and economic gap or unfairness, and make efforts to minimise discriminatory elements in all stages of algorithm development and use.

Transparency

Korea requires that an explanation is provided to users about the AI-enabled service's operation system and potential adverse-impacting factors used in AI enabled decisions.

Korea appears to be the only APAC jurisdiction that has explicitly set out that the duty of explanation over the telemarketing investment and insurance products must go beyond any explanation provided by a "chatbot", i.e. AI should not be used to explain AI. In addition, South Korea is the only jurisdiction intending to educate the population en-masse in relation to AI with its (non-binding) principles including a direction to propagate AI literacy population-wide by 2022

Data Quality

In Korea, the Ministry of the Interior and Safety has published a set of non-legally binding principles specifically reflecting data quality requirements, asking institutions to ensure fitness of data for intended service and apply pseudonymised or anonymised data as applicable.

[56] NITI (2019): https://niti.gov.in/sites/default/files/2019-01/NationalStrategy-for-AI-Discussion-Paper.pdf

Data Protection

In Korea, non-binding principles have been set out at a high-level that ask for lawful usage of personal data and protecting personal data and privacy while providing AI-enabled services.

Resilience

The Korea State Council in South Korea[57] recommended creating quality control standards to ensure AI reliability and stability to mitigate risks of resilience.

Cyber

The Korean State Council established an inter-ministry cooperation system to respond to new types of adverse effects of artificial intelligence use in 2020.The Korean State Council Principles on AI Initiative suggested institutions create quality control standards to vet AI reliability and stability from 2020.

4.6 Japan

In Japan, the Integrated Innovation Strategy Promotion Council advanced a set of "Social Principles of Human-Centric AI" [58] (Social Principles).

Fairness

The Social Principles provide that the use of AI should not infringe upon fundamental human rights that are guaranteed by international standards. The aim of the human-centric social-principle is to encourage proper use of AI. This includes:

- utilising AI to assist humans;
- requiring appropriate stakeholders involved in the design, delivery and utilisation of AI to be responsible for any consequences; and
- creating a user-friendly system for people to enjoy the benefits of AI.

Governance and Accountability

The principle of accountability has been adopted in Japan by the Integrated Innovation Strategy Promotion Council. in its "Social Principles of Human-Centric AI". The 2021 "AI Governance in Japan Ver 1.0 (interim report)" is also in consultation and looks to a goal-based governance suggested by the Study Group on a New Governance Model in Society 5.0 (Governance Model Study Group). It suggests regulatory intervention should be proportionate to the impact of risks for AI governance. This AI governance framework aims to be aligned globally.[59]

Data Protection

In Japan, the main theme from its financial regulator and privacy regulator with respect to data privacy is anonymisation. Both the Financial Services Agency and the PPC promote customer protection in terms of privacy and anonymity. Some of the measures recommended by the regulators include providing appropriate disclosures on handling anonymously processed information, ensuring

[57] Korea State Council (2019): http://www.korea.kr/news/pressReleaseView.do?newsId=156366736 (in Korean)
[58] Integrated Innovation Strategy Promotion Council (2019): https://www8.cao.go.jp/cstp/stmain/aisocialprinciples.pdf
[59] METI (2021): https://www.meti.go.jp/press/2020/01/20210115003/20210115003-3.pdf

security control measures and prohibiting identification of individual data subjects. The PPC submitted a Bill to Amend the Act on the Protection of Personal Information introducing new, data innovation-friendly concept of "Pseudonymously Processed Information", which requires the publication of its purposes of use.

Model Governance

Japan introduced a set of principles for AI utilisation that covers broader model governance [60]. The Japanese Ministry of Internal Affairs and Communications indicated that it is imperative to constantly review the Guidelines and revise them as necessary through international discussions, considering the extent of the progress of AI networking, because AI-related technologies and AI utilisation are expected to continue to advance dramatically.

Cyber

The Financial Services Agency of Japan adopted its "Financial Digitization Strategy" and encourages financial institutions to consider new and effective measures to mitigate and manage cyber-risks, promote international cooperation regarding cyber security and to consider measures to address new risks for the financial system associated with the development of digitisation.[61]

4.7 Thailand

Thailand's Ministry of Digital Economy and Society (MDES) Ministry has drafted its first AI ethics guideline in 2019. The final version of the guideline is yet to be released.

Fairness

The draft MDES AI ethics guidelines require AI solutions to consider diversity, minimise discrimination and bias and be able to prove fairness.

Data Quality

The MDES ethics guidelines recommend organisations to conduct quality control and data integrity checks as part of the development process of AI applications.

Resilience

The MDES ethics guidelines recommend creating quality control standards to ensure AI reliability and stability to mitigate risks of resilience.

Cyber

The MDES ethics guidelines encourage cooperation with the international community to mitigate against AI activity for improper motives and to leverage security principles from international leading practices.

[60] Japan the Conference toward AI Network Society (2019): https://www.soumu.go.jp/main_content/000658284.pdf
[61] FSA (2018): https://www.fsa.go.jp/en/news/2018/20180926/Financial_Services_Policy2018.pdf

4.8 Australia

Fairness

In Australia, the Department of Industry (DoI) released in November 2019 a framework of eight AI ethics principles. Fairness is one of them and the DoI states that throughout their lifecycle, AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.[62]

Governance and Accountability

The principle of accountability has been adopted in Australia in the DoI AI Ethics Principles which state that those responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

Data Quality

In Australia, the Office of the Australian Information Commissioner published a "Guide to data analytics and the Australian Privacy Principles"[63], which builds on the privacy principles in the specific context of data analytics and requires organisations to take rigorous steps to maintain quality of information used for data analytics.

Data Protection

The above "Guide to data analytics and the Australian Privacy Principles" also requires organisations to be careful with sensitive information. The guide extensively provides relevant key concepts when considering data analytics and privacy, and outlines how the Australian Privacy Principles apply to data analytics. One principle that appears to be adopted only in Australia is that of conducting a "privacy impact assessment" for data analytics projects. Australia also draws out the need to carefully consider whether uses are within the intended purpose.

Model Governance

The DoI AI Ethics Principles state that AI systems should be monitored and tested to ensure they continue to meet their intended purpose, and any identified problems should be addressed with ongoing risk management as appropriate. Responsibility should be clearly and appropriately identified, for ensuring that an AI system is robust and safe.

Resilience

As mentioned under the "Model Governance", the DoI also recommended that AI systems should be monitored and tested to ensure their resilience.

[62] DoI (2019): https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles
[63] Office of the Australian Information Commissioner (2018): https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-data-analytics-and-the-australian-privacy-principles/

4.9 EU

In the EU, the EC High Level expert group issued in April 2019 its "Ethics Guidelines for trustworthy AI[64](EU AI Ethics Guidelines). In addition, while we do not think there is a need for specific AI regulation for FS firms, we also note that the European Commission in April 2021 issued a "Proposal for a regulation laying down harmonised rules on AI"[65] (EU Proposed Regulation)

Fairness

The EU AI Ethics Guidelines state that unfair bias must be avoided, as it could have multiple negative implications, from the marginalisation of vulnerable groups, to the exacerbation of prejudice and discrimination.

Governance and Accountability

The EU AI Ethics Guidelines state that mechanisms should be put in place to ensure responsibility and accountability. There is a principle around accountability which includes requirements around auditability, minimisation and reporting of negative impacts, trade-offs and redress.

Transparency

The EU AI Guidelines state that the data, system and AI business models should be transparent and that traceability mechanisms can help achieve this. Moreover, AI systems and their decisions should be explained in a manner adapted to the stakeholder concerned. Humans need to be aware that they are interacting with an AI system, and must be informed of the system's capabilities and limitations.

Data Quality and Data Protection

The EU AI Ethics Guidelines state that besides ensuring full respect for privacy and data protection, adequate data governance mechanisms must also be ensured, taking into account the quality and integrity of the data, and ensuring legitimised access to data.

Moreover, the Proposed Regulation requires all training, validation and testing data sets to be subject to appropriate data governance and management practices, including relevant data preparation processing operations, such as annotation, labelling, cleaning, enrichment and aggregation. The Proposed Regulation prescribes for all training, validation and testing data sets to be subject to appropriate data governance and management practices, which requires that users of high-risk AI systems shall carry out a data protection impact assessment.

Model Governance

The Proposed Regulation requires human oversight throughout the AI system's lifecycle, which will enable individuals to understand capacities and limitations, fully monitor its operation and remain aware of tendency of automatically relying on output of high-risk AI.

Resilience

The EU AI Guidelines state that AI systems need to be resilient and secure. They need to be safe, ensuring a fall-back plan in case something goes wrong, as well as being accurate, reliable and reproducible.

The Proposed Regulation states that high-risk AI applications are subject to obligations around robustness and security.

Cybersecurity

The Proposal for a Regulation on a European Approach for Artificial Intelligence requires that high-risk AI systems be designed and developed in a way to achieve appropriate cybersecurity, and technical solutions for cybersecurity that are appropriate to the relevant circumstances and risks.

[64] EC (2019): https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[65] EC (2021): https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

4.10 UK

Fairness

The UK has actively focused on AI ethics by encouraging firms to build a robust data ethics framework. In 2018 and 2019, the UK Government published its Guidance[66] on Data Ethics Framework. This provided detailed guidelines on data protection law and the appropriate use of new technologies, with a holistic approach incorporating good practices in computing techniques, ethics and information assurance. They have also issued Guidance on Understanding AI Ethics and Safety which introduced AI ethics and provided a high-level overview of the ethical building blocks needed for the responsible delivery of an AI project. The ethical building blocks include building a culture of responsible innovation and a governance architecture to bring values and principles of ethical, fair, and safe AI together.

Separately, financial regulators, the BOE and FCA have followed and jointly published a survey[67] covering ethical issues on ML, which highlighted that ML is increasingly being used in UK financial sector. The key message from the survey was that AI systems must be designed and implemented in ways that incorporate measures to safeguard against potential risks of AI applications, such data quality issues (including biased data). Firms should identify ways to support safe, beneficial, ethical and resilient deployment of the technology across the UK financial sector, as well as understanding its impact on the wider economy. The BoE and FCA have stated they will consider repeating the survey in 2020.

Data Protection

The UK has explicitly considered individual's rights with respect to their personal data in AI systems. In the Information Commissioner's Office (ICO) Guidance on AI and data protection, the ICO identifies that an AI-specific area that should be covered includes individuals' rights to be forgotten, data portability and right to access personal data.

Governance and Accountability

The UK ICO in its guidance on AI and data protection[68] stresses that demonstrating embedding data protection by design and default into an organisation's culture and processes is an important element of accountability. The guidance also stresses the need for senior management involvement and accountability.

Also in the UK, the Office for Artificial Intelligence and the Alan Turing Institute have together recommended a Process Based Governance Framework. The Framework (i) clarifies the relevant team members and roles involved in each governance action, (ii) clarifies the relevant stages of the workflow in which intervention and targeted consideration are necessary to meet governance goals, (iii) provides explicit timeframes for any evaluations, follow-up actions, re-assessments, and continuous monitoring, and (iv) provides clear and well-defined protocols for logging activity and for implementing mechanisms to support end-to-end auditability.[69]

Transparency

The FCA in the UK has emphasised the following requirements in respect of the transparency and explainability principles of AI models:

- Customers should know when and how machines are involved in making decisions, whether it is about them or on their behalf;

[66] UK Government (2020): https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework#introduction
[67] FCA (2019): https://www.fca.org.uk/publication/research/research-note-on-machine-learning-in-uk-financial-services.pdf
[68] ICO: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-artificial-intelligence-and-data-protection/
[69] UK Government (2020): https://www.gov.uk/guidance/understanding-artificial-intelligence-ethics-and-safety and The Alan Turing Institute (2019): https://www.turing.ac.uk/sites/default/files/2019-06/understanding_artificial_intelligence_ethics_and_safety.pdf

- Transparent provision of consent;
- Boards will have to set the approach and level of detail involved in transparency, and that in turn will reflect the values of their organisations;
- "Sufficient" explainability should be the ultimate target;
- Board members should take a hard line on what "sufficient" means for them, and also what it should mean for the consumer; and
- It is vital that board members do not let them themselves be seduced by a 'black box knows best' argument.

The Office for Artificial Intelligence and Government Digital Service in the UK also highlights that transparency to stakeholders covers both how and why a model performed the way it did in a specific context.

Interpretability of AI applications has also been considered at a high-level by the UK ICO in its explainability guidance and AI Auditing Framework.

Model Governance

A research note by the FCA emphasises that firms need to validate AI applications before and after deployment. The most common validation methods are outcome-focused monitoring and testing against benchmarks. However, many firms note that AI validation frameworks still need to evolve in line with the nature, scale and complexity of AI applications.[70]

In the Guide to Artificial Intelligence Ethics and Safety jointly published by the FCA and the Turing Institute mentions that continuous inspection and monitoring of the system, so that its behaviour can be better predicted and understood, is essential to effective risk management. Rigorous protocols of testing, validating, verifying, and monitoring the safety of the system and the execution of self-assessments at each stage of the workflow, in order to ensure alignment with the safety objectives of accuracy, reliability, security, and robustness.[71]

Cyber

The UK's ICO AI Auditing Framework (non-binding) outlined the importance of a security policy, namely outsourcing risks, and re-identification risks. It also details the importance of testing and verification challenges.[72] The guidance also states that when an institution buys or outsources an AI solution, the institution should make sure that controllers and processors are able to fulfil their data protection obligations, in order to ensure there are no security breaches. The framework also suggests the necessity of human reviews for non-fully automated decision making, to ensure that safety checks are continuously in place.

[70] FCA (2019): https://www.fca.org.uk/publication/research/research-note-on-machine-learning-in-uk-financial-services.pdf
[71] The Alan Turing Institute (2019): https://www.turing.ac.uk/sites/default/files/2019-06/understanding_artificial_intelligence_ethics_and_safety.pdf
[72] ICO (2019): https://ico.org.uk/about-the-ico/news-and-events/ai-blog-ai-auditing-framework-call-for-input-final-considerations-and-next-steps

4.11 US

The CFTC issued via LabCFTC a "Primer on Artificial Intelligence in Financial Markets[73]" . In the primer, the CFTC outlines that to build effective AI systems, careful consideration must be given to the choice of algorithms, the sourcing of data, and the evolution of AI models, appropriate governance and controls are vital for AI to succeed, and that AI systems assist and augment, but cannot replace, human judgment.

CFTC also said that for long-term success, AI systems must be reliable, resilient, and trustworthy.

4.12 OECD AI Principles and Recommendations

Governance and Accountability

The OECD Recommendations on AI[74] provides the first intergovernmental standard for AI policies and a foundation on which to conduct further analysis and develop tools to support governments in their implementation efforts. The Recommendations specifically recommended that organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the OECD AI Principles [75].The OECD identified governance and accountability as some of the complementary values-based principles for the responsible stewardship of trustworthy AI, that organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the OECD AI Principles.

Model Governance and Resilience

The Recommendation sets out principles for the "responsible stewardship of trustworthy AI". In particular, the principles stated that the AI systems should be robust, secure and safe throughout their entire lifecycle so that they function appropriately and do not pose unreasonable safety risk. Although the recommendation has not given a definition for the measure of robustness, security and safety, we take this means minimising design risk and ensuring the performance of the AI model.

The Recommendation further calls on financial institutions to act and apply a systematic risk management approach to each phrase of the AI system lifecycle on a continuous basis to address the risks related to AI systems, including privacy, digital security, safety and bias.

Cyber

The OECD Principles on AI recommends that AI systems must function in a robust, secure and safe way throughout their life cycles and potential risks should be continually assessed and managed

[73] CFTC: LabCFTC Fintech Primers | CFTC
[74] OECD (2020): https://www.soumu.go.jp/main_content/000642218.pdf
[75] OECD: https://www.oecd.org/going-digital/ai/principles/

5 Glossary of Acronyms

This section provides a list of the Acronyms used in the report for your convenience.

ABS	Association of Banks in Singapore
AFMA	Australian Financial Markets Association
AI	Artificial Intelligence
AIDA	Artificial Intelligence for Data Analytics
AIMA	Alternative Investment Management Association
AML	Anti-Money Laundering
APAC	Asia Pacific
APRA	Australian Prudential Regulation Authority
ASIC	Australian Securities and Investments Commission
ASIFMA	Asia Securities Industry & Financial Markets Association
BAAI	Beijing Academy of Artificial Intelligence
BCBS	Basel Committee on Banking Supervision
BIS	Bank for International Settlements
BOE	Bank of England
DoI	Department of Industry (Australia)
EU	European Union
FCA	Financial Conduct Authority (UK)
FEAT	Fairness, Ethics, Accountability and Transparency
FRS	Federal Reserve System
FSB	Financial Stability Board
GDPR	General Data Protection Rules (EU)
GFIN	Global Financial Innovation Network
HKAB	Hong Kong Association of Banks
HKMA	Hong Kong Monetary Authority

ICO	Information Commissioner's Office (UK)
IOSCO	International Organization Of Securities Commissions
MAS	Monetary Authority of Singapore
METI	Ministry of Economy, Trade and Industry (Japan)
MDES	Ministry of Digital Economy and Society (Thailand)
ML	Machine Learning
MOST	Ministry of Science & Technology (China)
NEDO	New Energy and Industrial Technology Development Organisation (Japan)
Niti	National Institution for Transforming India
OCC	Office of the Controller of the Currency
OECD	Organisation for Economic Co-operation and Development
PCPD	Office of the Privacy Commissioner for Personal Data (Hong Kong)
PDPC	Personal Data Protection Commission (Singapore)
PPC	Personal Information Protection Commission (Japan)
PRA	Prudential Regulation Authority (UK)
SFC	Securities & Futures Commission (Hong Kong)
TPRM	Third-party risk management
UK	United Kingdom
US	United States of America