15 March 2024

To:
AI Verify Foundation and Infocomm Media Development Authority
Submitted by email to: info@aiverify.sg

ASIFMA Response to Proposed Model AI Governance Framework for Generative AI

Dear Sir/Madam,

On behalf of the Asia Securities and Financial Markets Association ("ASIFMA")[1] members, we thank you for the opportunity to respond to the *Proposed Model AI Governance Framework for Generative AI* developed by the AI Verify Foundation ("**AIVF**") and Infocomm Media Development Authority ("**IMDA**") (such draft framework being the "**Framework**").

The feedback set out in this response has been collected from ASIFMA's Fintech Working Group and AI Sub-Working Group, which have been closely following global, regional and local developments related to artificial intelligence ("**AI**") and emerging technologies in recent years.

ASIFMA has published its perspectives and recommendations regarding regulatory frameworks for AI in financial services in:

- our 2021 paper on "Enabling an Efficient Regulatory Environment for AI"[2]; and
- our recent addendum (in January 2024) to the above paper, on "Practical Considerations for Generative AI"[3].

The latter assesses the new and incremental challenges of generative AI; explores how these are already addressed by existing regulations, tools and governance frameworks; explores mitigants to such challenges; identifies gaps in current regulatory frameworks; and make suggestions on how to address the gaps with the aim to ensure the safe and responsible adoption of generative AI in the capital markets industry, so as

[1] ASIFMA is an independent, regional trade association with over 160 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 United States and AFME in Europe, ASIFMA also provides insights on global best practices and standards to benefit the region. More information about ASIFMA can be found at: www.asifma.org.

[2] ASIFMA (2021), Enabling an Efficient Regulatory Environment for AI, https://www.asifma.org/wp-content/uploads/2021/06/enabling-an-efficient-regulatory-environment-for-ai-report_june-2021.pdf

[3] ASIFMA (2024), Enabling an Efficient Regulatory Environment for AI – Practical Considerations for Generative AI, https://www.asifma.org/wp-content/uploads/2024/01/2024-asifma-gen-ai-paper-final-updated-08032024.pdf

to realise the full benefits of generative AI. We are encouraged to see that the views in our recent paper are generally aligned with that of the Framework.

General comments

We acknowledge the Framework's reference to the cloud industry's[4] shared responsibility models, which members believe is a step in the right direction. ASIFMA's 2024 Generative AI Paper also highlights how responsibility for the various controls will depend on the deployment model and use case of the relevant generative AI tool, and supports developers, deployers and the regulatory community collaborating to establish a mutual understanding of a shared responsibility framework. It is important to draw a distinction between shared responsibility and accountability. While financial institutions will remain accountable, responsibility for various controls will depend on the deployment model and use case of the relevant generative AI tool and use case.

Our members are generally supportive of the draft Framework, and have the following general suggestions for it:

- The Framework should clarify that its purpose is to serve as a best practice guidance for readers to develop their own risk management frameworks.
- The Framework should specify, for any relevant institution, the benefits of following the Framework and the extent to which the principles outlined in the Framework may be tailored for the relevant institution's specific situation or needs. The draft Framework currently does not seem to make that clear, other than stating that it "seeks to set forth a systematic and balanced approach to address generative AI concerns while continuing to facilitate innovation".

In addition, please find below some detailed comments on some sections of the draft Framework.

Sections of Framework

1. Accountability

We appreciate the reference to the cloud shared responsibility models, and have the following additional comments:

1. The cloud shared responsibility model is a good place to start, but it is important to adapt it to reflect the key differences between cloud infrastructure / platforms / software and generative AI. For example, if a foundation model is being relied on for Important Business Services[5], and that foundational model is then blocked or shut down[6], this could cause resilience risks (due to an inability to easily substitute a proprietary AI system – which is unlike cloud, where a potential fallback option is to switch to secondary data centres located elsewhere etc).

[4] This includes Google Cloud, Microsoft Azure and Amazon Web Services.

[5] It is defined by the FCA as a service provided by a *firm*, or by another *person* on behalf of the *firm*, to one or more *clients* of the *firm* which, if disrupted, could: (1) cause intolerable levels of harm to any one or more of the *firm's clients*; or (2) pose a risk to the soundness, stability or resilience of the *UK financial system* or the orderly operation of the financial markets.

[6] For example – ChatGPT was (between late March to late April 2023) blocked by OpenAI in Italy, after the Italian Data Protection Authority accused ChatGPT of unlawfully collecting users' data and failing to prevent underage users from accessing inappropriate material.

The shared responsibility model should be backed by, at a minimum, sufficient contractual provisions that in turn are backed by appropriate regulation. Such contractual provisions and regulations should also recognise the difficulties of negotiations with AI developers and deployers individually. This needs to be addressed by policymakers and organisations and should be included in the Framework. Concepts around indemnity and insurance do not adequately address the concerns of regulated entities like banks that face consumers as well as regulators and are held accountable by both parties, particularly as indemnity and insurance only come into consideration after a problem has occurred and does not address issues throughout the AI value chain such as transparency. It also increases costs, which reduces the viability of generative AI driven solutions. Further, while it is helpful that some generative AI providers have indemnified consumers against copyright issues in the generated output, it remains to be seen whether this is enforceable.

2. We also suggest that this section should encompass accountability of the models' impact – i.e. that it should verify compliance with substantive normative principles.

2. Data

Regarding the "Data" section in the Framework – we recommend including the potential for reidentification of individuals as a key risk within the Framework, given foundational models may link data from a variety of sources. Even if personally identifiable information is not explicitly used for training, a combination of characteristics from different sources could narrow down the search to a limited number of individuals, facilitating their identification. We also acknowledge that such risk is already referenced in a number of previous statements / documents from the Personal Data Protection Commission of Singapore – including:

- Guide to Basic Anonymisation (March 2022).
- Advisory Guidelines on the use of Personal Data in AI Recommendation and Decision Systems (March 2024).

The call to expand the available pool of trusted data is a welcome one. We recommend that the curation of a repository of representative data sets should be performed by independent organisations and governmental organisations working with local communities.

3. Trusted Development and Deployment

The need to balance transparency requirements with safeguarding proprietary information is a valid point. However, this is a source of tension between providers of generative AI foundational models and regulated entities, and it is difficult to strike a balance consistently across providers and across jurisdictions. Potential solutions could include independent audits, appropriate non-disclosures and non-compete agreements. We recommend that the Framework references such potential solutions and their potential role in balancing the interests of technology providers and regulated entities, including financial institutions.

4. Incident reporting

The "Incident Reporting" section of the Framework should consider the nature of an AI incident, which often plays out very differently from a traditional software or technology incident. AI-related incidents are not time-bound events as they usually are with Cloud or traditional software (e.g. infra failures or software bugs). For example, a case of systematic unjust bias may take weeks to confirm, due to the varying nature/seasonal fluctuations of demand and supply for the product/service under scrutiny. We recommend that the Framework identify patterns for defining/identifying AI incidents and assign responsibility for reporting and resolving them based on the nature of incident, especially where third-party providers are involved.

The Framework must carefully consider key regulatory developments across the world as AI regulation is in constant flux at this stage, and we urge regulatory harmonisation, especially with regard to reporting trigger events and timelines. We are supportive of the Framework's proposal that reporting should be proportionate and balanced between comprehensive reporting and practicality. Reporting requirements should be tailored to provide the right level of information while taking into account data confidentiality and privacy. Given that the regulatory landscape on (generative) AI is nascent and fast evolving, we would recommend against making references to specific (draft) regulations such as e.g. the EU AI Act.

5. Testing and assurance

The Framework should consider the potential for generative AI providers to tune their systems specifically to improve benchmark scores and thereby providing a false sense of security to consumers. There have been incidents in the past where AI developers have provided misleading benchmark scores. Any development of benchmark scoring regimes will need to have clearly identifiable standards, and with results regularly audited.

6. Security

We recommend referencing the National Institute of Standards and Technology Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations in this section.[7]

9. AI for Public Good

It will be beneficial to the reader in understanding whether this is a general vision for the uses of AI or a critical element of the governance framework.

ASIFMA wishes to thank the AIVF and IMDA for the opportunity to share this feedback on the proposed Framework. Members are supportive of continued dialogue between the AIVF, the IMDA, and the industry as regulatory standards and guidelines are being developed to ensure the appropriate calibration of the twin objectives of effectively managing risk whilst supporting responsible and sustainable innovation.

We welcome the opportunity to contribute to future consultations and remain at your disposal for further engagement or any further questions you might have. Please do not hesitate to reach out to us at lvanderloo@asifma.org or phone: +65 6622 5972.

Sincerely,

Laurence Van der Loo
Managing Director, Head of Technology & Operations
Asia Securities Industry & Financial Markets Association

[7] NIST (2024), Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations, https://csrc.nist.gov/pubs/ai/100/2/e2023/final