



## **Singapore launches AI Verify Foundation to shape the future of international AI standards through collaboration**

*IMDA has also published a discussion paper to share Singapore's practical and accretive approach to Generative AI governance*

**Singapore, 7 June 2023** – At the ATxAI conference, a part of Asia Tech x Singapore (ATxSG), **Mrs Josephine Teo, Singapore's Minister for Communications and Information** announced the launch of the AI Verify Foundation to harness the collective power and contributions of the global open source community to develop AI testing tools for the responsible use of AI. The Foundation will look to boost AI testing capabilities and assurance to meet the needs of companies and regulators globally. Seven pioneering premier members - the Infocomm Media Development Authority (IMDA), Aicadium (Temasek's AI Centre of Excellence), IBM, Microsoft, Google, Red Hat and Salesforce will guide the strategic directions and development of AI Verify roadmap. As a start, the Foundation will have also have more than 60 general members such as Adobe, DBS, Meta, SenseTime and Singapore Airlines<sup>1</sup>.

### **Building the Foundation for Trustworthy AI**

The launch of AI Verify Foundation will support the development and use of AI Verify to address risks of AI. AI Verify is an AI governance testing framework and software toolkit. first developed by IMDA in consultation with companies from different sectors and different scales. The Foundation will help to foster an open-source community to contribute to AI testing frameworks, code base, standards and best practices and create a neutral platform for open collaboration and idea-sharing on testing and governing AI.

Launched as a minimum viable product for international pilot last year, AI Verify attracted the interest of over 50 local and multinational companies including IBM, Dell, Hitachi and UBS. AI Verify is now available to the open source community and will benefit the global community by providing a testing framework and toolkit that is consistent with internationally recognised AI governance principles, e.g., those from EU, OECD, and Singapore. The AI Verify toolkit provides an integrated interface to generate testing reports that covers different governance principles for an AI system. It enables companies to be more transparent about their AI by sharing these reports with their stakeholders.

More information on the AI Verify Foundation and AI Verify toolkit is available in [Annex A](#). Collaborators will be able to get information on how they can join the foundation as well as access the open-source code at <https://aiverifyfoundation.sg>. Read about what our premium members have to say in [Annex B](#).

### **The Need for Collective Effort to Advance AI Testing Globally**

IMDA built AI Verify to help organisations objectively demonstrate responsible AI through standardised tests. However, AI testing technologies albeit growing, are still nascent. There is a need to crowd in the best expertise from across industry and research community to develop this area.

While recognising the potential risks of AI, Minister Teo noted that the government cannot do it alone. "The private sector with their expertise can participate meaningfully to achieve these goals with us," she added. She also assured that that amidst very real fears and concerns about AI's development, we will need to actively steer AI towards beneficial uses and away from bad ones. "This is core to how Singapore thinks about AI."

---

<sup>1</sup> The full list of members is available at <https://aiverifyfoundation.sg/foundation-members/>



## Discussions on a Global Scale

AI experts from around the world will be tackling pertinent issues on generative AI and AI governance. Some highlights include the distinguished panel of **Blaise Aguera y Arcas**, Vice President, Research, Google Research Int'l, US, **Kathy Baxter**, Principal Architect, Ethical AI Practice, Salesforce, US, **Shuicheng Yan**, Visiting Chief Scientist, BAAI, **Ben Brooks**, Head of Public Policy, Stability AI, **Michael Sellitto**, Head of Policy, Anthropic, and **Michael Zeller**, Head of AI Strategy and Solutions, Temasek, who will deep dive into key governance and alignment issues of generative AI that require multidisciplinary effort to solve pressing and long-term problems. The panel will also discuss approaches to close governance and technical gaps, with recommended areas for research and development.

Policymakers are also in for a treat as the panel on AI governance around the world, which features speakers such as **Kay Firth-Butterfield**, ED, Centre for Trustworthy Technology, WEF, **Ansgar Koene**, Global AI Ethics and Regulatory Leader, EY, **Elham Tabassi**, Associate Director for Emerging Technology, Information Technology Laboratory, National Institute of Standards and Technology and **Yi Zeng**, Professor and Director, International Research Center for AI Ethics and Governance, Institute of Automation, Chinese Academy of Sciences, will discuss how policymakers, industry and research communities globally must work together to address AI challenges faced collectively by humanity. They will also dissect different AI governance approaches around the world, taking a critical lens on their relevance in light of generative AI.

## Pathways for Policymakers for Generative AI

IMDA and Aicadium have published a discussion paper to share Singapore's approach to building an ecosystem for trusted and responsible adoption of Generative AI, in a way that spurs innovation and tap on its opportunities. The paper seeks to enhance discourse and invite like-minded partners from around the world to work with Singapore to ensure that Generative AI can be trusted and used in a safe and responsible manner.

The paper identifies six key risks that have emerged from Generative AI as well as a systems approach to enable a trusted and vibrant ecosystem. The approach provides an initial framework for policy makers to (i) strengthen the foundation of AI governance provided by existing frameworks to address the unique characteristic of Generative AI, (ii) address immediate concerns and (iii) invest for longer-term governance outcomes. The specific ideas (e.g. shared responsibility framework, disclosure standards) also seek to foster global interoperability, regardless of whether they are adopted as hard or soft laws. More information on the discussion paper is available in [Annex C](#).

The full Discussion Paper can be downloaded [here](#).

For details and the latest agenda on ATxSG, please visit: [www.asiatechxsg.com](http://www.asiatechxsg.com)

-END-



### **About the Infocomm Media Development Authority (IMDA)**

The Infocomm Media Development Authority (IMDA) leads Singapore's digital transformation by developing a vibrant digital economy and an inclusive digital society. As Architects of Singapore's Digital Future, we foster growth in Infocomm Technology and Media sectors in concert with progressive regulations, harness frontier technologies, and develop local talent and digital infrastructure ecosystems to establish Singapore as a digital metropolis.

### **About Asia Tech x Singapore (ATxSG)**

ATxSG 2023 is Asia's leading technology event jointly organised by Infocomm Media Development Authority (IMDA) and Informa Tech - supported by the Singapore Tourism Board (STB). The event consists of two main segments, ATxSummit and ATxEnterprise.

ATxSummit (6 – 7 June), the apex event of ATxSG held at Capella Singapore, comprises an invitation-only Plenary covering themes like generative AI, web 3.0 and trust, soonicorns and sustainability across four key pillars: Tech x Trust, Tech x Good, Tech x Builders and Tech x Creative. ATxSummit also features the ATxAI and SG Women in Tech conferences, alongside G2G and G2B closed-door roundtables to facilitate a closer partnership between the public sector and digital industry.

ATxEnterprise (7 – 9 June), organised by Informa Tech and held at Singapore Expo, will host conferences as well as exhibition marketplaces comprising B2B enterprises across Technology, Media, Infocomm, Satellite industries and start-ups. ATxEnterprise consists of BroadcastAsia, CommunicAsia, SatelliteAsia, TechXLR8 and InnovFest x Elevating Founders.

### **For media queries, please reach out to:**

Email: [Media@imda.gov.sg](mailto:Media@imda.gov.sg); [Sumita Sreedharan@imda.gov.sg](mailto:Sumita_Sreedharan@imda.gov.sg); [Yuen Hui Tian@imda.gov.sg](mailto:Yuen_Hui_Tian@imda.gov.sg)

### **Hill+Knowlton Strategies**

Email: [HKATxSG@hkstrategies.com](mailto:HKATxSG@hkstrategies.com)



## Annex A

### Fact Sheet - OPEN-SOURCING OF AI VERIFY AND SET UP OF AI VERIFY FOUNDATION

#### About AI Verify Foundation

IMDA has set up the AI Verify Foundation to harness the collective power and contributions of the global open source community to develop AI Verify testing tool for the responsible use of AI. The Foundation will boost AI testing capabilities and assurance to meet the needs of companies and regulators globally.

Current AI governance principles such as transparency, accountability, fairness, explainability, and robustness continue to apply to generative AI. The Foundation aims to crowd in expertise from the open-source community to expand AI Verify's capability to evaluate generative AI as sciences and technologies for AI testing develop.

The not-for-profit Foundation will:

- Foster a community to contribute to the use and development of AI testing frameworks, code base, standards, and best practices
- Create a neutral platform for open collaboration and idea-sharing on testing and governing AI
- Nurture a network of advocates for AI and drive broad adoption of AI testing through education and outreach

AI Verify Foundation has seven premier members, namely, Aicadium, Google, IBM, IMDA, Microsoft, Red Hat and Salesforce, who will set strategic directions and development roadmap of AI Verify. The Foundation also has more than 60 general members. For full list of members, please visit <https://aiverifyfoundation.sg/foundation-members/>

#### About AI Verify

IMDA had developed AI Verify, an AI Governance Testing Framework and Toolkit, to help organisations validate the performance of their AI systems against internationally recognised AI governance principles through standardised tests.

AI Verify is extensible so that additional toolkits (e.g. sector-specific governance frameworks) can be built on top of it. Contributors are encouraged to build components as plugins to AI Verify, and participate in growing the AI testing ecosystem.

Jurisdictions around the world coalesced around a set of key principles and requirements for trustworthy AI. These are aligned with AI Verify's testing framework which comprises 11 AI governance principles, namely,

- Transparency
- Explainability
- Repeatability/reproducibility
- Safety
- Security
- Robustness
- Fairness
- Data Governance
- Accountability
- Human agency and oversight



- Inclusive growth, social and environmental well-being

The testing processes comprises technical tests on three principles, namely, Fairness, Explainability, and Robustness. Process checks are applied to all 11 principles. The testing framework is consistent with internationally recognised AI governance principles, such as those from EU, OECD and Singapore.

AI Verify is a single integrated software toolkit that operates within the user's enterprise environment. It enables users to conduct technical tests on their AI models and record process checks. The toolkit then generates testing reports for the AI model under test. User companies can be more transparent about their AI by sharing these testing reports with their shareholders.

AI Verify can currently perform technical tests on common supervised-learning classification and regression models for most tabular and image datasets. AI Verify does not set ethical standards, neither does it guarantee AI systems tested will be completely safe or be free from risks or biases.

AI Verify was first developed in consultation with companies from different sectors and of different scales. These companies include AWS, DBS, Google, Meta, Microsoft, Singapore Airlines, NCS/LTA, Standard Chartered, UCARE.AI and XOPA. AI Verify was subsequently released in May 2022 for an international pilot, which attracted the interests of over 50 local and multinational companies, including Dell, Hitachi, IBM, and UBS.

## Annex B

### What the Premier Members of AI Verify Foundation have to say

<p><b>Aicadium (Temasek's AI Centre of Excellence)</b></p>	<p><i>Aicadium is proud to be a premier member of the AI Verify Foundation. With the rapid growth of Artificial Intelligence in business, government, and the daily lives of people, it is vitally important that AI is robust, fair, and safe.</i></p> <p><i>We look forward to working with the Foundation to take AI governance to the next level. We are committed to the development of rigorous, technical algorithmic audits and third party AI test lab capabilities, which we believe are an essential component of the AI ecosystem to help organisations deliver AI as a benefit to all.</i></p> <p><b>– Liu Feng-Yuan, Vice President, Business Development</b></p>
<p><b>Google</b></p>	<p><i>The AI Verify Foundation is an important step in ensuring that AI is used for good and that its benefits are shared by everyone. Its work will help to promote transparency and accountability in AI development, and to ensure that AI systems are fair, unbiased, and safe.</i></p> <p><i>Google is proud to support the Foundation in its mission and work collectively to advance AI's transformative potential in a bold and responsible way.</i></p> <p><b>- Michaela Browning, VP Government Affairs &amp; Public Policy, APAC</b></p>
<p><b>IBM</b></p>	<p><i>IBM believes in the potential of AI to transform businesses and society. Widespread adoption of AI in business can be, however, achieved by responsible and ethical development of systems to deliver the accurate and trusted outcomes. AI Verify is a significant step towards the implementation of AI governance, and IBM fully supports this framework in our AI governance platform.</i></p> <p><i>Organizations that want to use AI have a fundamental responsibility to foster trust in AI solutions, and IBM is dedicated to contributing and incorporating AI governance alongside the development of our responsible AI technology.</i></p> <p><b>- Colin Tan, IBM Singapore General Manager and Technology Leader</b></p>
<p><b>Microsoft</b></p>	<p><i>To foster trust in AI and ensure that its benefits are broadly distributed, we must commit to responsible practices around its development and deployment.</i></p> <p><i>We at Microsoft applaud the Singapore Government's leadership in this area. By creating practical resources like the AI Governance Testing Framework and Toolkit, Singapore is helping organizations build robust governance and testing processes. Everyone has a stake in this issue, and</i></p>

	<p><i>we are all better off when AI tools uphold respect for fairness, safety, and other fundamental rights.</i></p> <p>- <b>Brad Smith, President and Vice Chair</b></p>
<b>Red Hat</b>	<p><i>Red Hat is honored to join AI Verify Foundation in Singapore as a member. Increasingly, we see more and more powerful machine learning algorithms and generative AI tools being introduced into our society. It is inevitable that AI-driven solutions will transform how we live and work.</i></p> <p><i>By fostering the use of open source and community-driven technology, Red Hat believes it can contribute to the adoption of AI testing frameworks that are transparent, responsibly developed and trusted by the community we live in.</i></p> <p>- <b>Guna Chellappan, General Manager (Red Hat Singapore)</b></p>
<b>Salesforce</b>	<p><i>We are proud to be a part of Singapore’s leading efforts to advance responsible AI through partnerships between the public and private sector. At Salesforce, we believe that it is simply not enough to deliver only the technological capabilities of AI. We must also ensure that AI is safe and inclusive for all, by collaborating with and contributing to efforts led by governments, the industry, and society.</i></p> <p><i>The open-source nature of the AI Verify Foundation will create a level playing field for all, and allow leaders of responsible AI practices to contribute resources and share learnings for everyone to develop and deploy AI responsibly and safely. This will help build public trust in AI and lay the groundwork for an AI-led future.</i></p> <p>- <b>Kathy Baxter, Principal Architect, Ethical AI Practice</b></p>

## Annex C

### Fact Sheet - DISCUSSION PAPER ON GENERATIVE AI – IMPLICATIONS FOR TRUST AND GOVERNANCE

**Generative AI is uncovering a myriad of use-cases and opportunities** that are reshaping industries, revolutionising sectors and driving innovation. **At the same time, concerns have emerged**; from the risk of AI making gaffes to worries that it will take over the world. Amidst global discussions on Generative AI, IMDA, together with Aicadium, have co-written a discussion paper to share Singapore's approach and ideas for building an ecosystem for trusted and responsible adoption of Generative AI, in a way that spurs innovation and taps on its opportunities.

The paper considers various methods of assessing the risks of Generative AI and approaches towards AI Governance. It serves as a starting point for policy makers who wish to ensure that Generative AI is used in a safe and responsible manner, and that the most critical outcome – trust – is sustained.

The discussion paper is available [here](#).

#### *Overview of the Paper*

The paper identifies **six key risks** that have emerged from Generative AI - 1) mistakes and hallucinations, 2) privacy and confidentiality, 3) disinformation, toxicity and cyber-threats, 4) copyright challenges, 5) embedded bias and, 6) values and alignment.

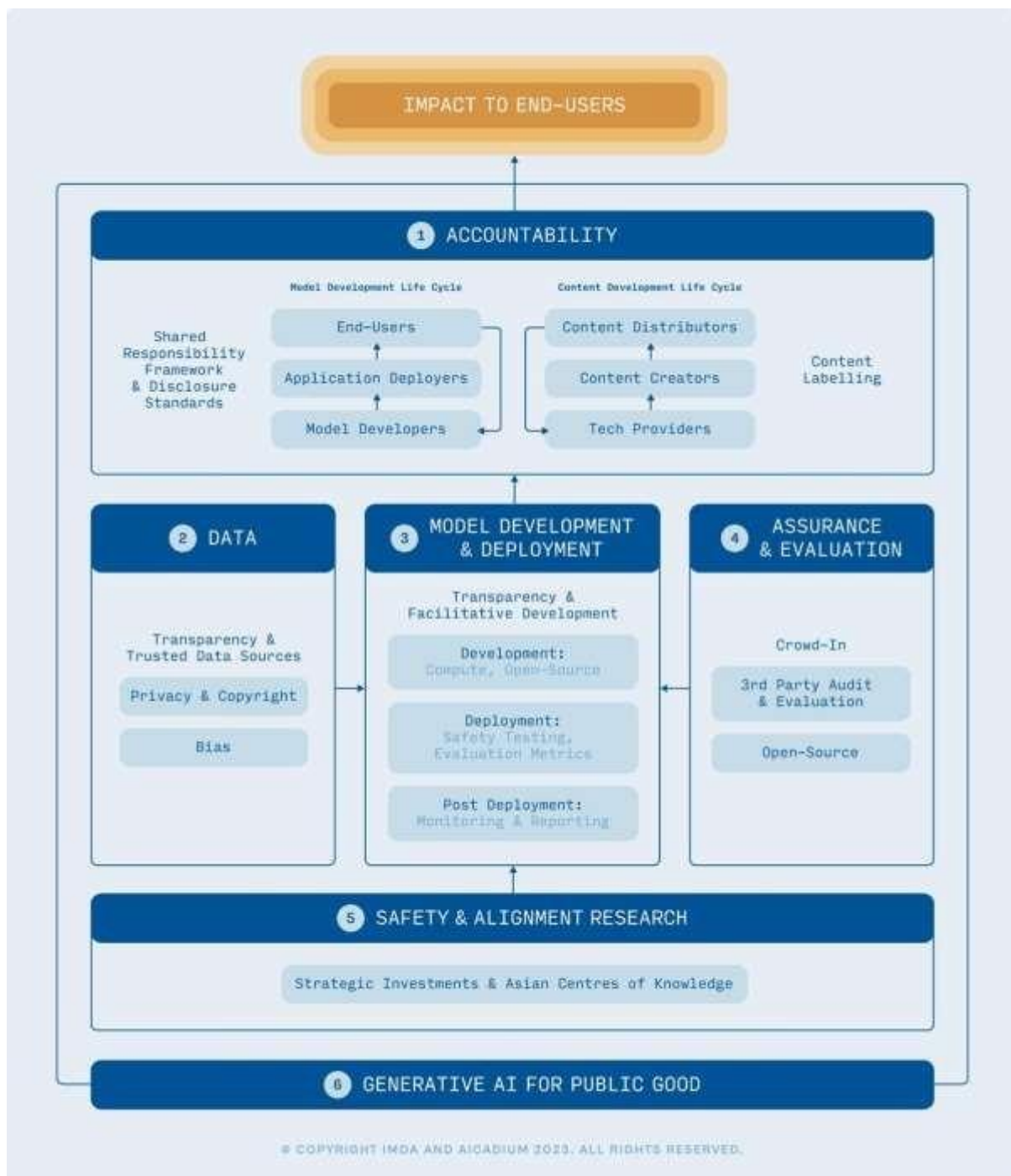
To address these challenges, Singapore adopts a **practical, risk-based and accretive** approach towards the governance of Generative AI by **building on existing AI governance principles**, such as those adopted by OECD, NIST and IMDA. Singapore's Model AI Governance Framework, for example, is based on the key governance principles - transparency, accountability, fairness, explainability, and robustness. While these principles and practices are applicable regardless of the types of AI deployed, policy adaptations will, nevertheless, be needed to consider the unique characteristics of Generative AI:

- Generative AI will increasingly form the **foundation** upon which other models/applications are built and there are concerns that problems inherent in these models could lead to wider issues. Governance frameworks will have to provide guidance on accountability between parties and across the development lifecycle, as well as address safety concerns in model development and deployment.
- These models are **generative** it may be increasingly difficult to distinguish AI-generated content - and people may become more susceptible to misinformation and online harms. As AI potentially surpasses human capacity at some levels, there are also concerns around how to control and align AI models.

The Discussion Paper identifies six dimensions that should be looked at in totality to enable a trusted and vibrant ecosystem of Generative AI. They are (1) **Accountability** as the basis for governance; Critical components of the model lifecycle from (2) **data** to (3) **model development and deployment** to (4) 3<sup>rd</sup> party **assurance and evaluation**; Longer term (5) **safety and alignment**; and (6) **Generative AI for public good** so that no one is left behind.

Collectively, the six dimensions provides an initial framework for policy makers to (i) strengthen the foundation of AI governance provided by existing frameworks to address the unique characteristic of Generative AI, (ii) address immediate concerns and (iii) invest for longer-term governance outcomes.

The specific ideas (e.g. shared responsibility framework, disclosure standards) also seek to foster **global interoperability**, regardless of whether they are adopted as hard or soft laws.



### **Considerations for Regulation**

There are practical considerations regarding the implementation and effectiveness of AI regulations. Technical tools, standards and technology to support regulatory implementation need to be ready before regulation can be effective. These are mostly still under development today.



Amidst the pressure to regulate, it is useful for governments to consider whether existing laws, such as sectoral legislation and data protection laws, can be tapped on and updated if necessary. Strongly interventionist regulations should also be carefully considered to tread the balance between risk mitigation and market innovation. For example, overly restrictive regulation on open-source models can stifle innovation by hindering collaboration and access. **Careful deliberation and a calibrated approach should be taken, while investing in capabilities and development of standards and tools.**

Singapore aspires to be a global leader in harnessing AI for the public good. We support the responsible development and deployment of AI so that its benefits may be enjoyed in a trusted and safe manner. If not properly governed, there is a risk that AI could be used for harms such as scams, cyberattacks, and misinformation.

We will continue to keep on top of the development of AI and will make **timely introduction and updates of targeted measures to uphold trust and safety in digital developments.** While Singapore does not currently intend to implement general AI regulation, the discussion paper itself is an example of how Singapore has taken concrete action to develop technical tools, standards and technology, which in turn lays the groundwork for clear and effective regulation in the future.

-END-

#### **About Aicadium**

Aicadium is a global technology company delivering AI-powered industrial computer vision products into the hands of enterprises. With offices in Singapore and San Diego, California, and an international team of data scientists, engineers, and business strategists, Aicadium is operationalizing AI within organizations where machine learning innovations were previously out of reach. As Temasek's AI Centre of Excellence, Aicadium identifies and develops advanced AI technologies, including areas of AI governance, regulation, and the ecosystem developments around AI assurance.