

Technology, Media & Telecommunications

First of its Kind Generative AI Evaluation Sandbox for Trusted AI

Introduction

Trust is at the core of most, if not all, successful ventures, initiatives and relationships. This human concept extends to the digital realm, and the advancement of artificial intelligence ("AI") solutions into an ever-increasing number of everyday tasks brings the issue of trusted AI models clearly into the limelight.

Seeking to ensure that generative AI is utilised in a safe and responsible manner where trust is sustained, the Infocomm Media Development Authority ("**IMDA**") and the AI Verify Foundation [announced](#), on 31 October 2023, the first of its kind Generative AI Evaluation Sandbox ("**Sandbox**"). The Sandbox will bring together key global players to build capabilities in the testing and evaluation of generative AI and is part of efforts to have a common standard approach to assess generative AI. The Sandbox utilises a new draft Evaluation Catalogue that sets out common baseline methods and recommendations for Large Language Models ("**LLMs**").

In this Update, we briefly explore the risks associated with generative AI models, the key aspects of the Sandbox and how Rajah & Tann can help you successfully navigate the myriad of issues relating to the adoption of AI solutions to meet your business needs.

Current Concerns Regarding LLMs

Generative AI refers to AI models that learn the underlying distribution of the data they are searching and then generate new content from that learned distribution. ChatGPT is an example of an extremely popular generative AI model.

Generative AI is being utilised in industries ranging from fashion to healthcare. However, our understanding of the risks of generative AI is still evolving. In the discussion paper "[Generative AI: Implications for Trust and Governance](#)" ("**Discussion Paper**"), IMDA detailed the key risks and harms of LLMs and proposed ideas for senior leaders in government and businesses on building an ecosystem for the trusted and responsible adoption of generative AI.

Key risks posed by generative AI highlighted in the Discussion Paper include:

- mistakes and "hallucinations";
- privacy and confidentiality;
- scaling disinformation, toxicity and cyber-threats;



Client Update: Singapore

2023 NOVEMBER

Technology, Media & Telecommunications

- copyright issues;
- ethical issues such as embedded biases which echo in downstream applications; and
- values, alignment and the difficulty of good instructions.

Endeavouring to address some of these risks, by working towards a common standard approach to assess generative AI and support the safe and trustworthy adoption of generative AI, IMDA is inviting industry partners to collaboratively build evaluation tools and capabilities in the Sandbox.

Key Aspects of the Sandbox

The Sandbox will utilise a new [Evaluation Catalogue](#) ("**Catalogue**"), as a shared resource, that details common baseline methods and recommendations for LLMs. The AI Verify Foundation welcomes initial comments and feedback on this draft Catalogue at info@aiverify.sg. To learn more about the AI Verify Foundation, please read our Legal Update regarding its launch [here](#).

Key model developers (including Google and Microsoft), application developers with concrete use cases (including DataRobot and OCBC) and third-party testers (including Deloitte and EY) have joined the Sandbox. The full list of participants is available [here](#). The Sandbox participants will assist in creating a more robust testing environment.

Key aspects of the Sandbox are:

- **Offering a common language for the evaluation of generative AI through the Evaluation Catalogue.** By offering a research-based categorisation of current evaluation benchmarks and methods, the Sandbox will provide a baseline for the evaluation of generative AI. The Evaluation Catalogue provides a base by compiling the existing commonly used technical testing tools and organising these tests according to what they test for and their methods, and recommending a baseline set of evaluation tests for use in generative AI products.
- **Creating a body of knowledge covering how generative AI products should be tested.** The Sandbox will help build evaluation capabilities beyond what currently resides with model developers. As the testing of generative AI should also include the application developers who build on top of the models, the Sandbox will involve players in the third-party testing ecosystem. This will enable model developers to understand what external testers require in responsible AI models. Where possible, each Sandbox use case should involve an upstream generative AI model developer, a downstream application deployer and a third-party tester to demonstrate how the different players in the ecosystem can collaborate. Moreover, by involving relevant regulators, such as the Singapore Personal Data Protection Commission (PDPC), the Sandbox allows for transparency regarding the needs of all parties along the supply chain.

Technology, Media & Telecommunications

- **Developing new benchmarks and tests.** The Sandbox use cases will likely reveal gaps in the current landscape of generative AI evaluations, particularly in currently underdeveloped domain-specific areas, such as human resources or security, and cultural-specific areas. The Sandbox will develop benchmarks for evaluating model performance in specific areas that are important for use cases, and for countries like Singapore that have cultural and language specificities.

Interested model and application developers, and third-party testers are invited by AI Verify Foundation and IMDA to participate in the Sandbox.

The Sandbox allows relevant parties to create trusted generative AI models that will have meaningful application in the Asian, as well as global, context. This development complements Singapore's evolving regulatory approach to AI. You may read about a recent proposed regulatory development in our Legal Update [here](#), which relates to data protection issues arising from the use of AI systems.

Concluding Words

The Sandbox is an important step in the journey towards building safe and trustworthy generative AI models. This exercise is reflective of Singapore's approach of engaging actively with AI stakeholders and providing a sandbox is the best way to deepen such sharing and learning through the work on the AI body of knowledge. This approach will also allow Singapore to develop protocols and standards with the AI community and international standards bodies, which is a necessary precursor to ensure that operational compliance can be achieved for any upcoming AI related laws and regulations in Singapore.

In this regard, the multi-disciplinary approach of our [Data and Digital Economy](#) model allows Rajah & Tann to prepare clients for the legal challenges of using generative AI models.

Apart from assisting clients in navigating regulatory and legal frameworks relevant to the use of AI models, our lawyers can also provide greater value-added legal services in collaboration with technical specialists in our [R&T Technologies](#) and [R&T Cybersecurity](#) team to offer multi-disciplinary and multi-legal solutions for clients.

Please feel free to contact our team below to discuss this development further.

Contacts

Technology, Media & Telecommunications



Rajesh Sreenivasan
Head, Technology, Media &
Telecommunications

T +65 6232 0751

rajesh@rajahtann.com



Steve Tan
Deputy Head, Technology,
Media & Telecommunications

T +65 6232 0786

steve.tan@rajahtann.com



Benjamin Cheong
Deputy Head, Technology, Media
& Telecommunications

T +65 6232 0738

benjamin.cheong@rajahtann.com

Click [here](#) for our Partners in Technology, Media and Telecommunications Practice.

Please feel free to also contact Knowledge Management at eOASIS@rajahtann.com

Regional Contacts

R&T SOK & HENG | *Cambodia*

R&T Sok & Heng Law Office

T +855 23 963 112 / 113

F +855 23 963 116

kh.rajahtannasia.com

RAJAH & TANN | *Myanmar*

Rajah & Tann Myanmar Company Limited

T +95 1 9345 343 / +95 1 9345 346

F +95 1 9345 348

mm.rajahtannasia.com

RAJAH & TANN 立杰上海

SHANGHAI REPRESENTATIVE OFFICE | *China*

Rajah & Tann Singapore LLP

Shanghai Representative Office

T +86 21 6120 8818

F +86 21 6120 8820

cn.rajahtannasia.com

GATMAYTAN YAP PATACSIL

GUTIERREZ & PROTACIO (C&G LAW) | *Philippines*

Gatmaytan Yap Patacsil Gutierrez & Protacio (C&G Law)

T +632 8894 0377 to 79 / +632 8894 4931 to 32

F +632 8552 1977 to 78

www.cagatlaw.com

ASSEGAF HAMZAH & PARTNERS | *Indonesia*

Assegaf Hamzah & Partners

Jakarta Office

T +62 21 2555 7800

F +62 21 2555 7899

Surabaya Office

T +62 31 5116 4550

F +62 31 5116 4560

www.ahp.co.id

RAJAH & TANN | *Singapore*

Rajah & Tann Singapore LLP

T +65 6535 3600

sg.rajahtannasia.com

RAJAH & TANN | *Thailand*

R&T Asia (Thailand) Limited

T +66 2 656 1991

F +66 2 656 0833

th.rajahtannasia.com

RAJAH & TANN | *Lao PDR*

Rajah & Tann (Laos) Co., Ltd.

T +856 21 454 239

F +856 21 285 261

la.rajahtannasia.com

RAJAH & TANN LCT LAWYERS | *Vietnam*

Rajah & Tann LCT Lawyers

Ho Chi Minh City Office

T +84 28 3821 2382 / +84 28 3821 2673

F +84 28 3520 8206

CHRISTOPHER & LEE ONG | *Malaysia*

Christopher & Lee Ong

T +60 3 2273 1919

F +60 3 2273 8310

www.christopherleeong.com

Hanoi Office

T +84 24 3267 6127

F +84 24 3267 6128

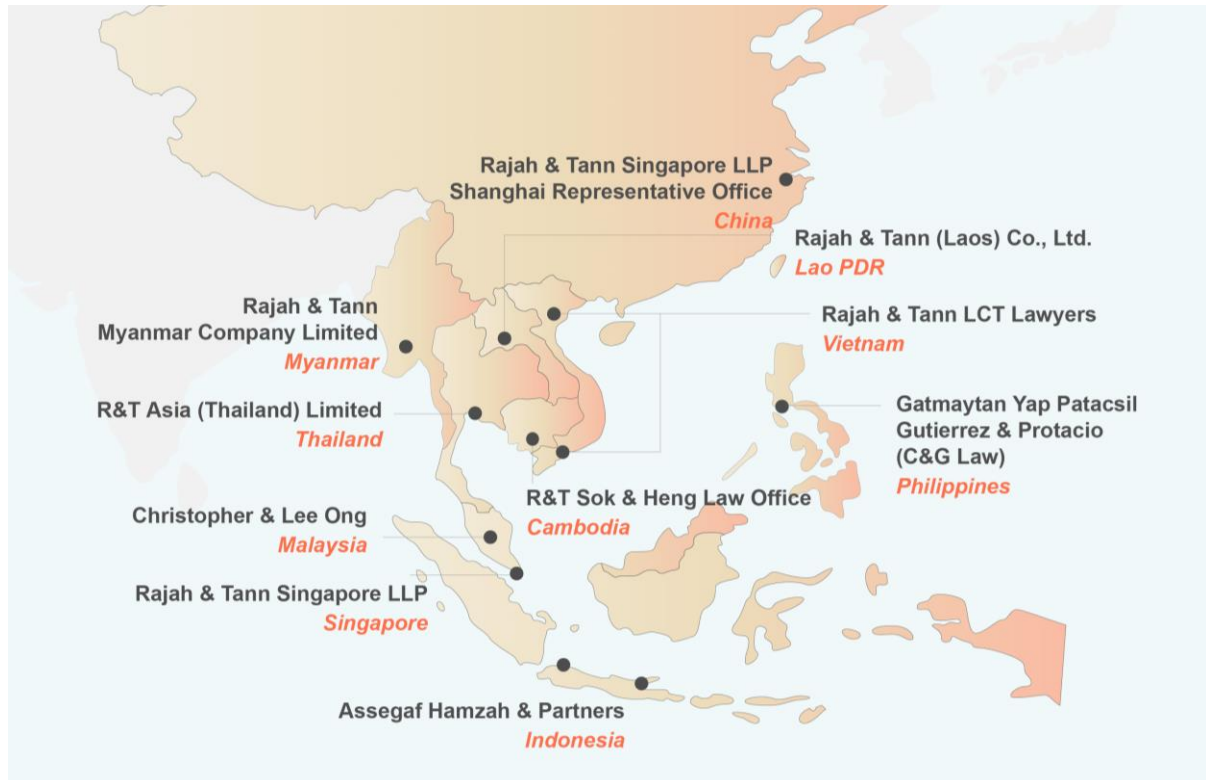
www.rajahtannlct.com

Rajah & Tann Asia is a network of legal practices based in Asia.

Member firms are independently constituted and regulated in accordance with relevant local legal requirements. Services provided by a member firm are governed by the terms of engagement between the member firm and the client.

This update is solely intended to provide general information and does not provide any advice or create any relationship, whether legally binding or otherwise. Rajah & Tann Asia and its member firms do not accept, and fully disclaim, responsibility for any loss or damage which may result from accessing or relying on this update.

Our Regional Presence



Rajah & Tann Singapore LLP is one of the largest full-service law firms in Singapore, providing high quality advice to an impressive list of clients. We place strong emphasis on promptness, accessibility and reliability in dealing with clients. At the same time, the firm strives towards a practical yet creative approach in dealing with business and commercial problems. As the Singapore member firm of the Lex Mundi Network, we are able to offer access to excellent legal expertise in more than 100 countries.

Rajah & Tann Singapore LLP is part of Rajah & Tann Asia, a network of local law firms in Cambodia, China, Indonesia, Lao PDR, Malaysia, Myanmar, the Philippines, Singapore, Thailand and Vietnam. Our Asian network also includes regional desks focused on Brunei, Japan and South Asia.

The contents of this Update are owned by Rajah & Tann Singapore LLP and subject to copyright protection under the laws of Singapore and, through international treaties, other countries. No part of this Update may be reproduced, licensed, sold, published, transmitted, modified, adapted, publicly displayed, broadcast (including storage in any medium by electronic means whether or not transiently for any purpose save as permitted herein) without the prior written permission of Rajah & Tann Singapore LLP.

Please note also that whilst the information in this Update is correct to the best of our knowledge and belief at the time of writing, it is only intended to provide a general guide to the subject matter and should not be treated as a substitute for specific professional advice for any particular course of action as such information may not suit your specific business and operational requirements. It is to your advantage to seek legal advice for your specific situation. In this regard, you may call the lawyer you normally deal with in Rajah & Tann Singapore LLP or email Knowledge Management at eOASIS@rajahtann.com.