



AI Verify Foundation accredits its first five AI testers as IMDA launches an AI agent testing kit

Description

The AI Verify Foundation has accredited its first five firms to test AI systems for safety and reliability: AIDX Tech, Knovel Engineering, PwC Singapore, Resaro and Vulcan. The accreditation was part of a package the Infocomm Media Development Authority (IMDA) announced on 14 October, which also included a testing kit for AI agents and a new taskforce on AI-driven cyber threats.

Organised by:



AI ASSURANCE EXCHANGE

AGENTIC AI EDITOR

NOVOTEL SINGAPORE ON



14 OCTOBER



IMDA announced the three measures at the AI Assurance Exchange at Novotel Singapore on Stevens on 14 October. Image: AI Verify Foundation

What the accreditation covers

The AI Tester Accreditation Programme, or AI TAP, assesses firms on two fronts: technical skill in testing generative AI systems, and readiness to sell testing services to AI developers and deployers.

Firms are accredited for specific testing methods, benchmarking and red teaming, across five risk areas: hallucination and inaccuracy, undesirable content, data leakage, vulnerability to adversarial prompts, and bias and fairness.

The five firms are not accredited for the same things. AIDX Tech is accredited for benchmarking and Vulcan for red teaming, while Knovel Engineering, PwC Singapore and Resaro hold both. The foundation says AI TAP recognises demonstrated competence only. It does not certify that an accredited firm's testing is independent, or that any AI system it tests is completely safe.

Organisations that run into problems with an accredited tester can send feedback to assurance@aiverify.sg.

A four-step kit for testing agents

IMDA's Starter Kit for Testing Agentic Systems for Safety and Reliability is aimed at organisations deploying AI agents, which can use tools, reach sensitive data and carry out tasks for a user. The kit sets out four steps: identify risks, design tests, analyse results and logs, and keep monitoring and retesting after launch, CNA reported. It draws on the Global AI Assurance Sandbox, the Singapore AI Safety Institute and consultations with more than 20 industry participants.

“When we give AI more agency, we need greater assurance over how that agency is exercised,” IMDA chief executive Ng Cher Pong said at the AI Assurance Exchange: Agentic AI Edition at Novotel Singapore on Stevens, AsiaOne reported.

A taskforce on AI-driven cyber threats

The Frontier AI Resilience and Security Taskforce, or FAIRS, brings together more than 20 leaders and experts from telecommunications, media, cybersecurity and AI. According to CNA, it will work on faster responses to AI-accelerated attacks, strengthening critical systems as AI gets better at finding weaknesses, and pooling intelligence on new attack techniques.

FAIRS is due to publish practical guidelines and best practices in the first half of 2027. The announcement came a day after Singapore launched a national cybersecurity strategy on 13 October.

Date Created

17/10/2026

Author

vanessakoh