

Singapore funds S\$70m programme to scale the SEA-LION regional language model

Source copy kept by wheresthe.ai for the story "Singapore funds S\$70m programme to scale the SEA-LION regional language model" (<https://wheresthe.ai/d/sg-national-multimodal-llm-programme-sea-lion-2023-12/>).

Source: <https://imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2023/sg-to-develop-southeast-asias-first-llm-ecosystem>

Text saved 18 Oct 2026 (the page text through a reader, because the site refused a direct fetch).

SHA-256 of the saved text: a3961f2be57aa3110d1a0aff12381a2f4aa702cce6926d0a0ef8ea0d72d8e1b7

Singapore's National Multimodal Large Language Model (LLM) Programme to drive regional and domestic innovation in AI, build the nation's capabilities in AI R&D and develop AI talent

****SINGAPORE – 04 DEC 2023****

1. Singapore's Infocomm Media Development Authority (IMDA) is partnering with AI Singapore (AISG) and the Agency for Science, Technology and Research (A*STAR) to launch the National Multimodal LLM Programme (NMLP). This effort will enhance the nation's capabilities in Artificial Intelligence (AI) research and innovation, leveraging high-performance computing (HPC) resources by the National Supercomputing Centre (NSCC). This programme marks a significant leap forward in building the next bounds of AI capabilities for Singapore and the region.

2. Funded by the National Research Foundation, Singapore (NRF), the new S\$70 million Initiative will develop Singapore's research and engineering capabilities in multi-modal Large Language Models. In line with the Research, Innovation and Enterprise (RIE) 2025 plan, the two-year national-level initiative will also support Singapore's National AI Strategy 2.0 (NAIS 2.0), which was launched by Deputy Prime Minister and Minister for Finance, Mr Lawrence Wong, at the SCAI. The NAIS 2.0 aims to nurture Singapore as a global leader in AI solutions by 2030 by developing and deploying scalable, impactful AI solutions in key sectors of high value and relevance to citizens and businesses. The National Multimodal LLM Programme will:

* **Build skilled AI talent** in Singapore by providing funding and access to high-end computing for local researchers and engineers. This effort will enable the development of a pool of skilled AI professionals to drive innovation and technological excellence in developing AI solutions in Singapore. * **Foster a thriving AI industry** to develop LLM-enabled solutions for greater productivity and new opportunities for businesses. This programme will foster collaborations with industry partners to provide a conducive environment to catalyse the development of novel AI enterprise use cases in a rapidly evolving landscape of AI technology. * **Enable Singapore to build a trusted environment** in using AI, as the programme will provide a deeper understanding of how LLM work and further research AI governance. The NMLP will **develop a base model with regional context** that can understand Singapore's and the region's unique linguistic characteristics and multi-lingual environment.

> This national effort underscores Singapore's commitment to become a global AI hub. Language is an essential enabler for collaboration. By investing in talent and investing in large language AI models for regional languages, we want to foster industry collaboration across borders and drive the next wave of AI innovation in Southeast Asia.

Dr Ong Chen Hui

Assistant Chief Executive, Biztech Group, IMDA

3. As technology evolves rapidly, there is a strategic need to develop sovereign capabilities in LLMs. Singapore and the region's local and regional cultures, values and norms differ from those of Western countries, where most large language models originate. A cornerstone of this initiative is the development of multimodal and localised LLMs for Singapore and the region to understand context and values related to the diverse cultures and languages of Southeast Asia, for example, managing context-switching between languages in multilingual Singapore.

A collaborative effort in shaping Singapore's AI ambitions

4. This effort in research, engineering, governance, computing, and innovative enterprise AI use cases will develop models and use cases within the next two years. To do this, we will build on early outcomes of AISG's

recently launched SEA-LION 1 (Southeast Asian Languages in One Network) model, an open-sourced large language model that is more representative of Southeast Asia's cultural contexts and linguistic nuances. SEA-LION is designed to be relatively smaller, flexible and faster than the commonly used LLMs in the market today. It represents a relatively inexpensive and efficient option for organisations, especially the many cost-sensitive and throughput-constrained enterprises in Southeast Asia, to incorporate AI into their workflows. This project will further develop SEA-LION to 30-50B parameters in size 2, and extend SEA-LION into a multimodal speech-text model.

The latter will draw on A*STAR's Institute for Infocomm Research's (I 2 R) work in speech and language research that has been applied widely in language transcription and translation supporting various agencies and companies in the private sector. I 2 R's multimodal speech-text foundation model could help identify non-verbal cues and enable SEA-LION to have a closer read of the user intent.

Developing trust with AI

5. IMDA's AI Governance Model Framework and AI Verify support tech companies and businesses to harness the collective power and contributions of the global open-source community to develop AI testing tools for the responsible use of AI. Establishing trust and necessary guardrails are ever more critical in the GenAI space to ensure that our model fulfils AI ethics and safety principles. IMDA will bring together the ecosystem in data, governance, testing and benchmarking alongside research in AI safety and alignment.

6. The development of the National Multimodal Large Language Model (LLM) Programme follows from initial efforts by various government agencies to leverage and expand Singapore's AI capabilities. These include the launch of AISG in 2017 as the national programme in AI, A*STAR's Centre for Frontier AI Research, as well as the Research Programme on the Governance of AI and Data Use undertaken by SMU's Centre of AI and Data Governance.

1 SEA-LION is a family of LLMs that is specifically pre-trained and instruct-tuned for the Southeast Asian (SEA) region (<https://github.com/aisingapore/sealion>)

2 Large Language Models are machine learning models that are trained to understand, generate human language. They are often measured in the number of parameters it contains, for example, GPT-4 (~1,700 billion parameters), Llama-2 model (70 billion parameters).

Stay tuned for our newsletter