ILMU System Card

Last updated: August 11, 2025

Contents

The **ILMU System Card** provides information about the ILMU foundation model, including its training and evaluations, known limitations, safety performance, and mitigation approaches. This document may be periodically updated to reflect the latest improvements or evaluations.

For more detailed technical information, please see the **ILMUchat Technical Report**.

ILMU - Model Overview

Name: ILMU (foundation model family)
Developers: YTL AI Labs
Release date: August 12, 2025
Modalities: Text (prompts and file uploads), images, and voice.

1. Introduction

ILMU is a family of Artificial Intelligence (AI) models – developed, owned, and operated entirely within Malaysia. As Malaysia's frontier foundation model, ILMU is fluent in Bahasa Melayu (BM) and deeply grounded in local needs, values, and context, ensuring its responses are culturally aware, contextually intelligent, and legally compliant in Malaysia.

ILMUchat, our consumer-facing product, delivers seamless text, voice, and vision interactions. For enterprise customers, the ILMU API extends these capabilities with advanced image and video generation. Every experience is powered by the ILMU foundation model, ensuring consistent quality, versatility, and cultural alignment.

To ensure responsible deployment, ILMU's voice, vision and text capabilities have undergone rigorous safety evaluations to identify and mitigate potential risks. Our assessments confirm that ILMU does not pose greater safety risks than current state-of-the-art AI systems.

2. Model Data

Data sources used in model training comprised public datasets, licensed content, proprietary collections (through our partnership with Universiti Malaya) and internally curated materials. The data sets were selected to reflect Malaysia's multilingual and multicultural society. All data were filtered for quality and pre-processed to remove potentially sensitive content and personally identifiable information.

Training Dataset

ILMU was pre-trained on a large dataset covering various domains and data formats, including publicly accessibly web content, programming code (in multiple languages), mathematical data, images, and audio.

Post-Training Dataset

ILMU was fine-tuned on a curated dataset explicitly tailored to the Malaysian context. This data set consisted of instruction-response pairs that reflect local human preferences and tool interactions, enabling ILMU to provide responses that are more helpful, culturally-sensitive and safe.

3. Intended Use and Limitations

Intended Usage

ILMU is optimized for use cases that require deep contextual understanding of Malaysian culture. ILMU is particularly versed in cultural references and BM proficiency. ILMU also

excels at interpreting sentences that combine multiple languages, such as BM and English ("Manglish").

Known Limitations

Like all foundation models, ILMU may occasionally exhibit limitations, including hallucinations, challenges with causal reasoning, complex logical deductions, and counterfactual reasoning. Additionally, while proficient in multiple languages, its strongest capabilities are specifically in the Malay language. Its knowledge cut-off date is May 2024.

For further details on our approaches to mitigate associated risks and harms, please refer to the **Ethics, Safety & Governance** section below.

4. Capabilities & Evaluations

We evaluate ILMU's capabilities across multiple dimensions of language understanding and cultural alignment using established multitask benchmarks. These include the widely used **MMLU** (Massive Multitask Language Understanding) for general knowledge, **CMMLU** for Chinese language proficiency, and **MalayMMLU** for Malay language and Malaysian cultural contexts.

As shown in Table 1, ILMU delivers competitive performance across all three benchmarks, including setting a new state-of-the-art on MalayMMLU. This performance underscores ILMU's versatility across diverse linguistic environments while maintaining strong local adaptation.

Model	MMLU	CMMLU	MalayMMLU
ILMU	80.39	**83.64**	**87.20**
GPT-4o	83.94	76.53	84.97
DeepSeek-V3	**84.15**	81.67	80.56
Llama-3.1-70B	82.32	72.04	78.07
SEA-LION-v3-70B	85.48	71.45	78.03
Sahabat-AI-v2-70B	84.09	72.03	78.31
Mallam 2.5 Small	65.80	31.31	71.35
Merdeka LLM HR 3B	43.84	59.91	57.28

Table 1: Comparison of ILMU and other models on Multitask Language Understanding (MLU) benchmarks. **Bold** indicates the best score per column.

4.1. MalayMMLU: Benchmarking Malay Language Proficiency

MalayMMLU is the first dedicated multitask benchmark for the Malay language, developed by local researchers to evaluate LLMs on both linguistic proficiency and cultural relevance. It contains 24,213 curriculum-aligned questions spanning primary (Year 1–6) and secondary

(Form 1–5) education levels in Malaysia, covering five domains and 22 subjects.

Table 2 presents ILMU's MalayMMLU results by subject area. ILMU ranks highest in all domains, demonstrating strong performance not only in language but also in STEM, humanities, and social sciences. This comprehensive capability reflects its close alignment with Malaysian linguistic norms and its suitability for culturally grounded applications.

Model	Language	STEM	Humanities	Social Science	Others	Overall
ILMU	89.36	88.05	88.40	85.44	85.08	87.20
GPT-4o	87.64	83.54	87.78	82.84	82.34	84.97
DeepSeek-v3	83.13	83.91	78.84	78.25	78.00	80.56
GPT-5	83.59	78.10	76.50	80.73	75.44	79.53
Sahabat-AI	80.60	78.31	81.25	76.39	74.93	78.31
Llama 3.1	79.44	78.76	81.00	76.50	75.10	78.07
SEALION	79.20	78.76	81.11	76.45	75.22	78.03
Mallam 2.5 Small (Mesolitica)	73.00	70.00	71.00	72.00	70.00	71.53
Merdeka-LLM (Agmo)	56.92	57.63	60.36	56.82	55.10	57.28
Falcon3-10B	54.77	58.20	60.17	56.76	54.04	56.38

Table 2: MalayMMLU benchmark scores by subject area for various models.

5. Ethics, Safety & Governance

5.1 Safety Approach

ILMU was developed in close collaboration with internal safety and responsibility teams. Internal and external red-teaming exercises were conducted to identify risks throughout the development lifecycle. Mitigation strategies have been implemented across all stages, from data curation and training to ongoing post-deployment monitoring.

Human-in-the-Loop Safety Oversight

We believe human oversight is integral to ensuring a culturally aligned, responsible, and safe model. As such, human-in-the-loop is a fundamental principle of ILMU's safety framework: human experts are involved at every critical stage of the model lifecycle.

- **Training Data**: Human experts curated and filtered training data to ensure content quality, cultural appropriateness, and removal of sensitive or harmful information.
- **Red Teaming**: Internal and external teams conducted exercises to surface safety vulnerabilities, edge cases, cultural misalignment and potential misuse across diverse domains.
- **Testing and Evaluation**: Human reviewers assessed model outputs, especially for safety-critical use cases, ensuring alignment with ethical, cultural, and legal standards in Malaysia.
- **Governance Review**: YTL AI Labs governance body reviewed all safety evaluations and approved the model prior to deployment.

5.2 Risk Detection & Monitoring

By using ILMU products, users agree not to engage in deceptive, misleading, or harmful activities or attempt to bypass built-in safeguards. YTL AI Labs actively monitors its services to enforce these policies and has mechanisms in place to address any violations detected across all ILMU-powered platforms.

To do so, we integrate real-time safety filtering and risk detection mechanisms. User input and model responses are evaluated against internal risk classifications. High-risk interactions are blocked or flagged for internal review. When safety thresholds are breached, the model is designed to refuse to respond. These safeguards are supported by continuous monitoring of user activity and behavior patterns, with processes in place for rapid response and future updates as new risks emerge.

5.3 Cultural Bias & Representation

Bias mitigation is deeply integrated into ILMU's development process, with a strong emphasis on cultural alignment and representational fairness in the Malaysian context. Evaluations included benchmarks specific to Malaysian culture and laws, assessing political sensitivities (e.g., royalty, race, religion), sexual orientation, gender identity, and minority representation, guided by principles in Malaysia's Federal Constitution.

While ILMU has been carefully developed to align with Malaysian values and use cases, certain design choices introduce the following limitations:

- **Malaysia-Specific Grounding**: ILMU is specifically optimized for Malaysian context, cultural norms, and majority perspectives. As a result, its accuracy and relevance may diminish when handling topics outside the Malaysian context. Furthermore, in thethe cases of conflicting information, the model is designed to prioritize authoritative local sources, potentially limiting its ability to reflect plural or global viewpoints.
- **Audio Robustness Constraints**: Although ILMU has been fine-tuned to accommodate Malaysia's multilingual and multiracial environment, its audio capabilities are tailored to a conversational Malaysian-English ("Manglish") tone. As such, unless prompted otherwise, ILMU defaults to speech that reflects local linguistic styles, including informal expressions and slang commonly found in "Manglish."

These limitations reflect conscious tradeoffs to ensure cultural alignment, but they are areas of active monitoring and potential future refinement.