



# Wholly Made in Malaysia by Malaysians for Malaysians

---

YTL AI Labs

## Abstract

This technical report introduces **ILMUchat**, the *first* Malaysian multimodal generative AI (GenAI) assistant that exhibits *text*, *vision*, and *audio* capabilities. ILMUchat is designed to address the needs of Malaysians and is developed with a deep understanding of the country’s multilingual and multicultural landscape. ILMUchat is built specifically for the Malaysian context: (i) *Text* - ILMUchat supports the major languages used in Malaysia, including Malay, Chinese, English, and Manglish, the distinctive local blend of English with Malaysian expressions; (ii) *Vision* - it interprets visual content relevant to Malaysian culture; and (iii) *Audio* - it accurately transcribes and outputs Malaysian-accented speech. These capabilities enable ILMUchat to effectively navigate Malaysia’s diverse linguistic and cultural environment. Our model achieves top performance on *MalayMMLU*, a multitask benchmark for the Malay language, and delivers strong results on English and Chinese benchmarks such as *MMLU* and *CMMLU*. For context on the broader research and development initiative, visit our official site: <https://www.ytlailabs.com/>

## 1 Introduction

Generative Artificial Intelligence (GenAI) assistants, such as ChatGPT [1], Gemini [6], and Qwen [7] have transformed global interaction by combining text, vision, and speech into unified user experiences. However, outside the innovation hubs of the United States and China, there is a noticeable gap in region-specific multimodal GenAI assistants. In Southeast Asia, initiatives like AI Singapore’s SEA-LION [4] offer valuable yet solely text-based solutions, supporting languages such as Indonesian, Vietnamese, and Filipino. Similarly, Malaysia’s Mesolitica provides MaLLaM [8], another text-only assistant focused on local linguistic contexts.

Recognizing this significant opportunity and the growing demand for culturally grounded AI experiences, YTL AI Labs proudly introduces **ILMUchat**, Malaysia’s first multimodal GenAI assistant developed specifically for Malaysians. *Wholly made in Malaysia by Malaysians for Malaysians*, ILMUchat integrates advanced capabilities in *text*, *vision*, and *speech*. It is thoughtfully designed to reflect nuances of Malaysian culture, values, and everyday contexts. ILMUchat is not merely a technological milestone. It marks a bold step forward in Malaysian innovation, strengthening connections between artificial intelligence (AI) and national identity.

## 2 Overview of ILMUchat

ILMUchat is a consumer-facing generative AI assistant for Malaysia. It serves as the national counterpart to ChatGPT, enabling natural interaction through *text*, *image*, and *speech*. These capabilities are

made possible by **ILMU**, the underlying foundation model that provides the linguistic intelligence and cultural grounding behind ILMUchat. As a result, ILMUchat can generate responses that are not only coherent and context-aware, but also deeply aligned with Malaysian language and culture.

ILMU is a foundation model, with its base model trained from scratch

## 2.1 Foundation Model: ILMU

ILMU is a uniquely Malaysian foundation model trained from scratch as a large language model. It is a **fully Malaysian intellectual property**, created by YTL AI Labs in partnership with Universiti Malaya, and supported by a team of local researchers and international advisors.

The name ILMU reflects four core pillars that guide its design and deployment:

- (i) **Intelek** (Intellect): Delivers insightful, context-aware responses
- (ii) **Luhur** (Integrity): Built with strong ethical safeguards and alignment to Malaysian values
- (iii) **Malaysia**: Understands and communicates naturally like a Malaysian
- (iv) **Untukmu** (For You): Designed for all Malaysians, across all backgrounds and sectors

The development of ILMU prioritises safety, transparency, and cultural alignment at every stage of the training process. Its pipeline was designed through close collaboration with both internal and external safety teams. Human experts were involved throughout the lifecycle, from dataset selection and filtering to model evaluation and deployment oversight, ensuring that all aspects of the system align with Malaysia’s legal frameworks, ethical standards, and cultural norms.

Building on this foundation, ILMU was trained from scratch using a carefully curated mix of publicly available data, licensed third-party corpora, and Malaysia-centric sources. The inclusion of local content ensures that ILMU reflects the country’s linguistic diversity and social realities. Guided by expert review, this training strategy balances broad language understanding with deep cultural grounding. Although the exact dataset composition and model parameter sizes are not publicly disclosed, ILMU has been trained across multiple model scales to support a wide range of downstream applications, including ILMUchat (as presented in this paper) and Ryt AI for Ryt Bank<sup>1</sup>, from lightweight edge deployments to high-capacity server inference.

## 2.2 Text Capability

ILMUchat is a multilingual GenAI assistant capable of effortlessly switching between Malay, English and Chinese, even within the same conversation. This makes it well-suited for Malaysia’s multicultural and multiethnic society, where these languages are frequently mixed in daily communication. As shown in Figure 1, ILMUchat can respond fluidly across languages, offering inclusive support to users from diverse linguistic and cultural backgrounds.

Beyond multilingualism, ILMUchat is also attuned to the linguistic diversity of the Malay language, including its standard, literary, colloquial, and royal variants. This reflects not only language ability but also sensitivity to Malaysia’s social and cultural contexts.

**Standard Malay** is the formal variant used in official communication. In contrast, literary Malay encompasses styles such as *hikayat*, a traditional form of prose or verse used to narrate historical or legendary tales, often with romanticized or epic elements. In Figure 2a, ILMUchat is prompted to generate a news article in the *hikayat* style, demonstrating its capability in classical literary expression.

**Colloquial Malay**, or Bahasa Pasar (literally “market language”), is widely used in informal settings. It includes shortened forms such as “*jap*” for *sekejap* (translated as *wait*), as well as code-switching with English words like *relax*, *chill*, or *best*. ILMUchat is able to engage in natural conversation using Bahasa Pasar, as illustrated in Figure 2b.

**Royal Malay**, or Bahasa Istana, is a refined form of Malay used in communication with the monarchy. As shown in Figure 2c, ILMUchat can accurately convert standard Malay sentences into Bahasa Istana, reflecting an understanding of the language’s sociolinguistic hierarchy.

---

<sup>1</sup><https://www.rytbank.my/>

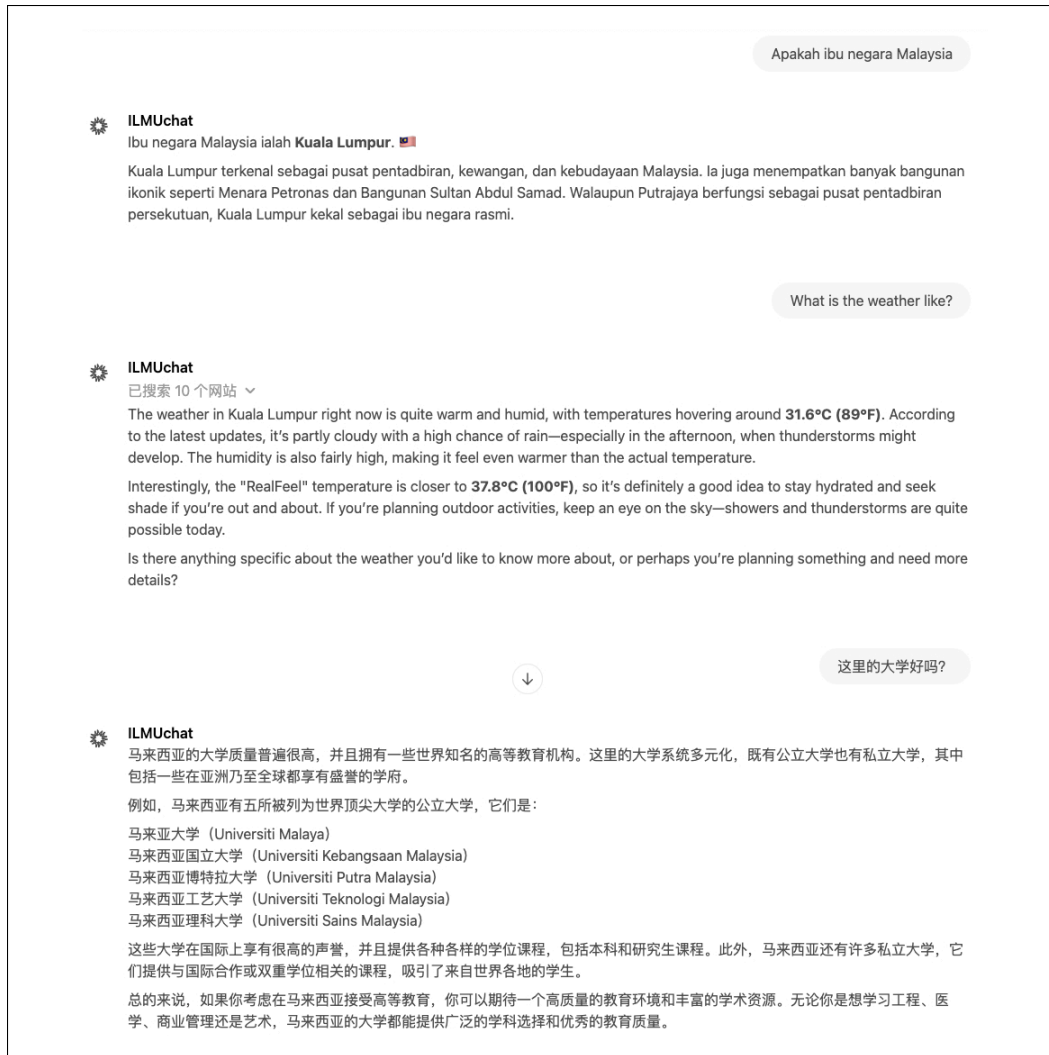


Figure 1: ILMUchat can converse in Malay, English and Chinese and interchanges between these languages seamlessly.

ILMUchat also demonstrates **strong general-purpose versatility** across a wide range of tasks, including objective question answering, mathematics, text generation, programming, subjective Q&A, logical reasoning, and text processing. The prompts are listed in Table 1, and the corresponding responses are shown in Figures 3 - 9, highlighting ILMUchat's broad capability in handling diverse text-based interactions.

Beyond general task performance, ILMUchat is also designed to be linguistically and factually aligned with the Malaysian context. In Figure 10, ILMUchat answers a question about the historical figure *Hang Tuah* in grammatically accurate Malay, while DeepSeek's response includes Indonesian words. Furthermore, Figure 11 illustrates ILMUchat's superior performance in answering factual questions specific to Malaysia, where other models such as DeepSeek and Gemini 2.5 Flash provide incorrect answers.

Table 1: Text capabilities of ILMUchat on different tasks and their corresponding prompts. The responses of ILMUchat to these prompts are shown in Figures 3 to 4.

Category	Task	Prompt
Objective QA	Scientific Knowledge	In Malay, briefly explain how monsoons affect Malaysia's weather patterns.
	Humanities & History	Who was the first Yang di-Pertuan Agong of Malaysia
	Practical Tools	Convert 170 cm to feet and inches
	Literature & Arts	Who is Usman Awang and why is he important to Malaysian literature?
Mathematics	Real-time Queries	What is the weather at Malacca at this moment?
	Elementary Math	Apakah 25% daripada 100?
Text Generation	Modern Math	Jelaskan teorem binomial
	Practical Writing Styles	tuliskan sebuah surat rasmi kepada Jabatan Kesihatan Cheras untuk laporkan terdapatnya jejentik dalam takungan air di taman permainan yang berdekatan dengan rumah aku.
	Stylized Text	tuliskan syair tentang kepentingan ilmu
	Summarization	Summarize RMK-13.
Code-related Instructions	Data Construction	Write 5 supervised fine-tuning dataset in Malay language about Nasi Lemak for LLM, return in JSONL of {"query": query text, "answer": LLM answer}.
	Code Generation	Write a Python script to convert Ringgit to Singapore Dollars using an API.
Subjective QA	Code Q&A	What's the difference between list and set in Python?
	Opinion & Evaluation	What's your opinion on Malaysia's policy on electric vehicles?
Logical Reasoning	Suggestion & Planning	Suggest a 3-day travel itinerary in Langkawi for a family with kids.
	Causal Reasoning	Why did the flash floods happen in Kuala Lumpur despite new drainage upgrades?
	Common Sense Reasoning	If it's 1PM during Ramadan, should you eat in public in Malaysia?
Text Processing	Editing & Proofreading	Proofread this sentence: 'He go to the pasar malam every Friday.'
	Translation	Translate this sentence to Malay: 'Education is the key to success.'
	Identification & Categorization	Classify this tweet as positive, negative, or neutral: 'Can't wait for the long weekend in Malaysia!'



(a) Literature

(b) Colloquial Malay

(c) Royal Malay

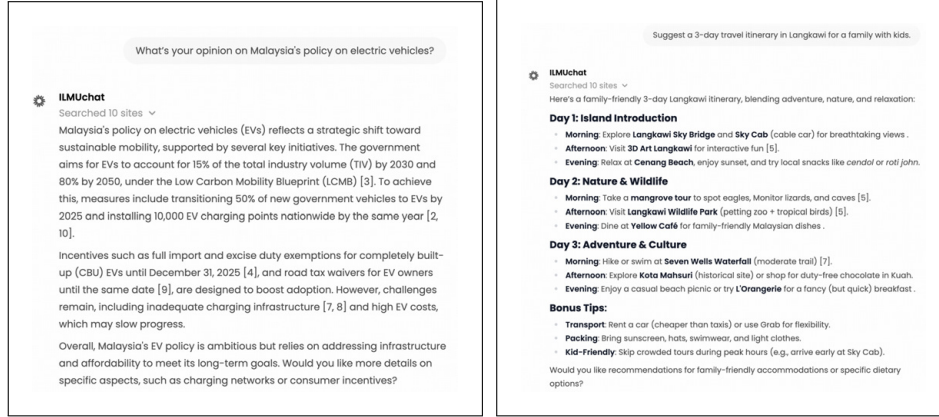
Figure 2: ILMUchat can generate responses in diverse Malay styles: (a) literary hikayat, (b) colloquial Bahasa Pasar, and (c) formal Bahasa Istana used in royal contexts.



(a) Editing & Proofreading

(b) Translation

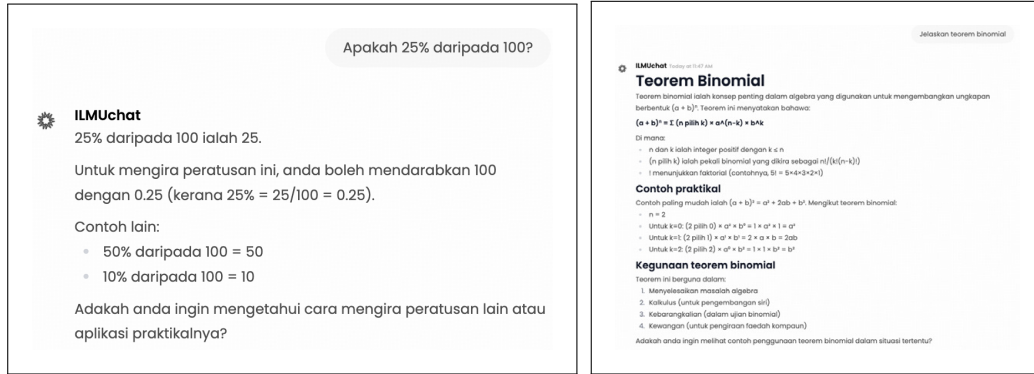
Figure 3: Example of queries and responses of ILMUchat for text processing



(a) Opinion & Evaluation

(b) Suggestion & Planning

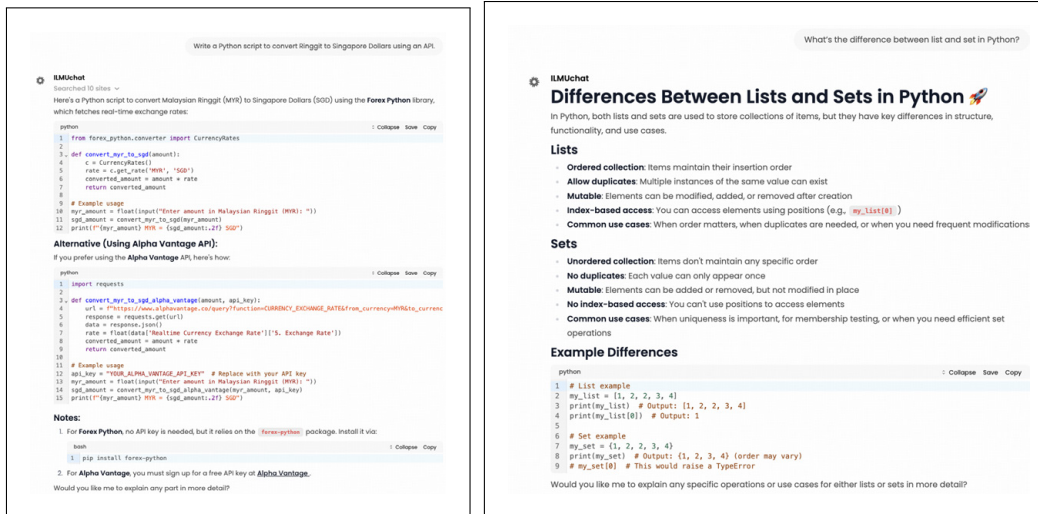
Figure 4: Example of queries and responses of ILMUchat for subjective question answering (QA)



(a) Elementary Mathematics

(b) Modern Mathematics

Figure 5: Examples of mathematical queries and ILMUchat's responses, demonstrating its ability to explain percentages and binomial theorem in grammatically accurate Malay.



(a) Code Generation

(b) Code Instruction

Figure 6: Examples of queries and ILMUchat's responses to code-related instructions, demonstrating its ability to understand prompts and generate syntactically correct, contextually relevant code.

In Malay, briefly explain how monsoons affect Malaysia's weather patterns.

**ILMUchat**

## Monsun dan Corak Cuaca di Malaysia

Monsun memainkan peranan penting dalam menentukan corak cuaca di Malaysia, yang mengalami iklim tropika lembap. Terdapat dua musim monsun utama:

- Monsun Timur Laut (November hingga Mac)** – Monsun ini membawa hujan lebat, terutamanya ke negeri-negeri pantai timur Semenanjung Malaysia (Kelantan, Terengganu, Pahang) dan barat Sarawak. Angin bertiup dari arah timur laut, menyebabkan laut bergelora dan risiko banjir di kawasan rendah.
- Monsun Barat Daya (Mei hingga September)** – Monsun ini biasanya lebih kering dengan angin bertiup dari arah barat daya. Hujan berkurangan, tetapi ribut petir dan hujan lebat boleh berlaku pada waktu petang di kawasan pantai barat.

Monsun juga mempengaruhi aktiviti pertanian, perikanan, dan pelancongan. Contohnya, nelayan sering mengelak turun ke laut semasa Monsun Timur Laut kerana ombak kuat.

Ingin tahu lebih lanjut tentang kesan monsun terhadap aktiviti harian?

Who is Usman Awang and why is he important to Malaysian literature?

**ILMUchat**

Searched 10 sites

Usman Awang was a highly celebrated Malaysian poet, playwright, and novelist, widely regarded as the 'people's poet' for his ability to resonate deeply with Malaysians across diverse backgrounds [1, 4]. Born on 12 July 1929, his literary contributions left an indelible mark on Malaysia's cultural landscape until his passing on 29 November 2001 [2, 4]. His importance stems from several key aspects:

- Poet Laureate Status:** Usman Awang was bestowed the prestigious title of Malaysian National Laureate (Sasterawan Negara) in 1983, solidifying his position as a cornerstone of national literature [5]. This honour recognised his profound impact on shaping Malaysia's literary identity.
- Versatile Writer:** While primarily known as a poet, he was also an accomplished short story writer and playwright, showcasing his multifaceted talent [6]. His works, such as the poem "Pak Uthi," provided sharp social commentary, reflecting his lifelong concern for marginalised communities [8, 9].
- Cultural Influence:** Usman Awang's literature transcended ethnic and age boundaries, winning numerous awards including the SEA Write Award. His writing inspired generations, with many crediting his works for fostering a love of Malay literature and achieving academic success [7].
- Legacy in Tourism:** Recent studies highlight his enduring relevance, positioning him as a significant figure in Malaysia's literary tourism, underscoring his continued cultural importance [10].

In summary, Usman Awang remains a pivotal figure for elevating Malaysia's literary voice, advocating for social awareness through art, and uniting communities through timeless storytelling. His legacy continues to inspire both scholars and the public alike.

(a) Scientific Knowledge

(b) Literature and Arts

Who was the first Yang di-Pertuan Agong of Malaysia

**ILMUchat**

The first Yang di-Pertuan Agong of Malaysia was Tuanku Abdul Rahman ibni Almarhum Tuanku Muhammad, who reigned from 31 August 1957 to 1 April 1960. He was instrumental in Malaysia's independence and played a pivotal role in shaping the nation's early governance. His leadership set the foundation for the constitutional monarchy system that continues to this day. Would you like to know more about his contributions or the significance of his reign?

What is the weather at Malacca at this moment ?

**ILMUchat** Today at 10:59 AM

Searched 10 sites

According to the most recent weather data, Malacca is currently experiencing cloudy skies with a low temperature of 78°F (approximately 26°C). The winds are light and variable, and the humidity is around 84%. There is a 9% chance of rain at the moment [1, 4, 5]. The UV index is at 0 out of 11, indicating minimal ultraviolet exposure. For the upcoming evening, expect scattered showers and thunderstorms with heavy downpours at times [3, 6].

Would you like more details about the forecast for the next few days?

Convert 170 cm to feet and inches

**ILMUchat**

170 centimetres is equal to approximately 5 feet 7 inches.

To break this down:

- 1 foot = 30.48 centimetres
- 170 ÷ 30.48 ≈ 5.58 feet
- The remaining 0.58 feet × 12 inches/foot ≈ 7 inches

So, 170 cm = 5 feet 7 inches.

Would you like this converted to any other units?

(c) Humanities and History

(d) Real-time Queries

(e) Practical Tools

Figure 7: Example of queries and responses of ILMUchat for objective question answering (QA).

Why did the flash floods happen in Kuala Lumpur despite new drainage upgrades?

**ILMUchat**

Searched 9 sites

Flash floods in Kuala Lumpur persist despite drainage upgrades due to a combination of factors, including rapid urbanisation, inadequate drainage systems, and environmental changes. Here's a breakdown:

- Urbanisation and Impervious Surfaces:** The city's expansion has replaced green spaces (which absorb rainwater) with concrete surfaces, increasing runoff and overwhelming drainage systems [5, 6]. This reduces the natural absorption capacity of the land.
- Inadequate Drainage Systems:** Even with upgrades, drainage systems may still be insufficient to handle intense rainfall. Clogged drains and poorly maintained infrastructure exacerbate the issue, causing water to back up quickly [4, 8].
- Rapid Rise in Water Levels:** Flash floods are characterised by a sudden surge in river and waterway levels, often exceeding capacity due to heavy rainfall [2]. This can overwhelm even recently upgraded systems.
- Environmental Factors:** Overdevelopment and concretisation disrupt natural water flow, while weather conditions (e.g., intense rain) further strain the system [7]. The Klang Valley's topography also contributes to recurrent flooding [3].
- Maintenance Gaps:** Upgrades alone aren't enough if drains aren't regularly cleared of debris, which can block water flow and lead to floods [8].

To mitigate this, a holistic approach is needed—balancing urban growth with sustainable drainage solutions, regular maintenance, and community involvement in keeping drains clear. Would you like to know more about specific drainage projects or flood prevention efforts in KL?

If it's IPM during Ramadan, should you eat in public in Malaysia?

**ILMUchat**

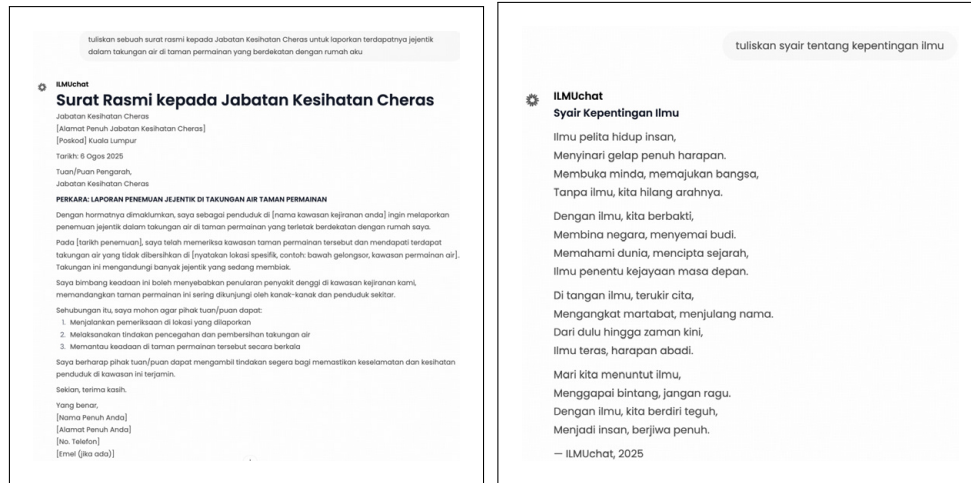
It is generally discouraged and respectful to avoid eating, drinking, or smoking in public during daylight hours in Malaysia during Ramadan, especially if you are a Muslim. This is because fasting from dawn to sunset is a central pillar of Ramadan for Muslims, and public observance is part of the cultural respect for the holy month. While non-Muslims are not obligated to fast, many choose to do so out of respect for their Muslim friends, colleagues, and the community. If you need to eat due to health reasons or other unavoidable circumstances, it is advisable to do so discreetly and respectfully, away from public view. If you are unsure, it's best to follow the practices of the majority and show consideration during this sacred period. Would you like to know more about Ramadan customs in Malaysia?

(a) Causal Reasoning

(b) Common sense Reasoning

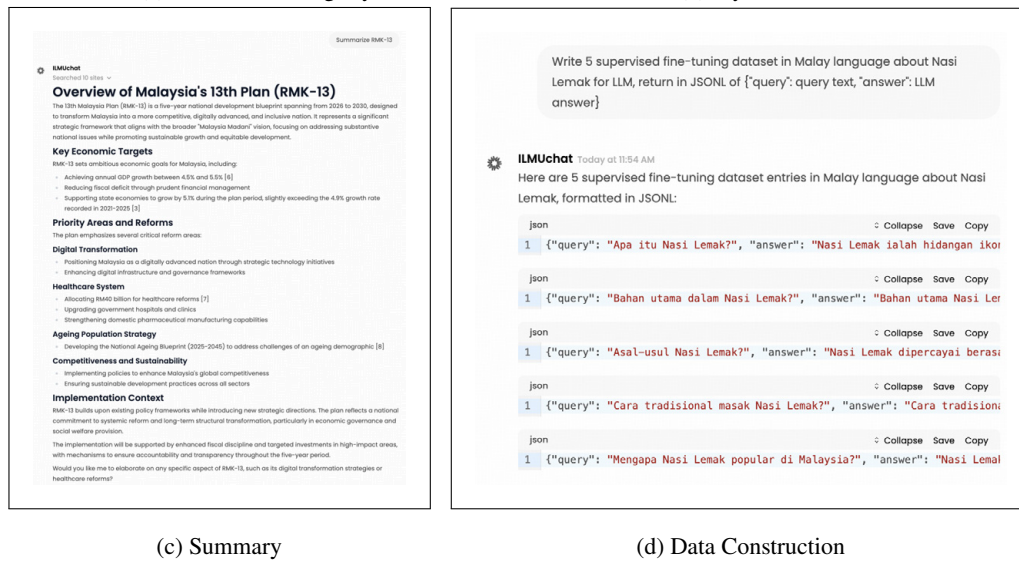
Figure 8: Example of queries and responses of ILMUchat for logical reasoning





(a) Practical Writing Styles

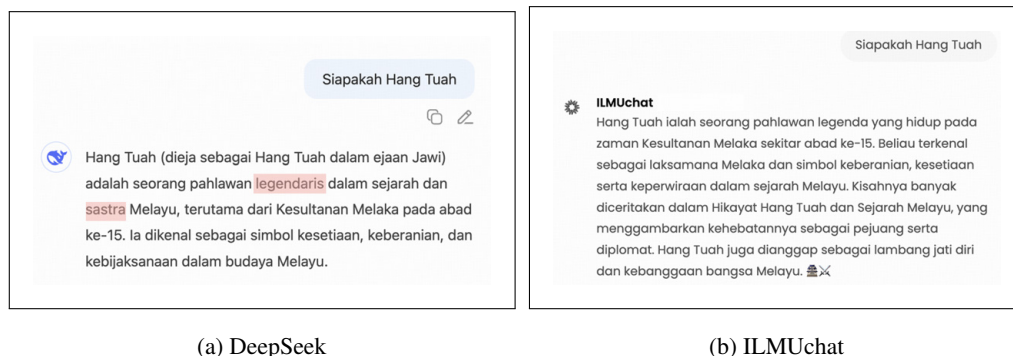
(b) Stylized Texts



(c) Summary

(d) Data Construction

Figure 9: Example of queries and responses of ILMUchat for text generation



(a) DeepSeek

(b) ILMUchat

Figure 10: Comparison of the responses on Malaysian historical figure - *Hang Tuah*. (a) DeepSeek's response contains Indonesian words (red), (b) ILMUchat produces grammatically correct Malay.

<p>Apakah nama PMX dalam bahasa cina?</p> <p>"PMX" dalam bahasa Mandarin biasanya merujuk kepada <b>Presiden Prancis Emmanuel Macron</b>, yang disebut 马克龙 (Mǎkèlóng).</p> <p>Namun, jika yang Anda maksud adalah sesuatu yang lain (misalnya singkatan khusus), mohon berikan konteks lebih lanjut agar saya bisa memberikan jawaban yang tepat! 🙏</p> <p>✗</p>	<p>Apakah nama PMX dalam bahasa cina?</p> <p>Tun Dr. Mahathir Mohamad, bekas Perdana Menteri Malaysia, dikenali dalam bahasa Cina sebagai 马哈蒂尔 (Mǎhātīěr).</p> <p>✗</p>	<p>Apakah nama PMX dalam bahasa cina?</p> <p><b>ILMUchat</b> Today at 2:44 PM Perdana Menteri Malaysia, Dato' Seri Anwar Ibrahim, sering dirujuk sebagai "PMX" dalam media dan perbualan tempatan. Dalam bahasa Cina, beliau dikenali sebagai 安华 (Ānhuá). Nama ini digunakan secara meluas dalam komuniti Cina di Malaysia serta media berbahasa Cina.</p> <p>✓</p>
<p>Wilayah Persekutuan yang kedua terbesar di Malaysia</p> <p>Wilayah Persekutuan yang kedua terbesar di Malaysia adalah <b>Putrajaya</b>.</p> <p>Berikut adalah senarai Wilayah Persekutuan di Malaysia mengikut saiz:</p> <ol style="list-style-type: none"> <li>1. <b>Kuala Lumpur</b> – 243 km²</li> <li>2. <b>Putrajaya</b> – 49 km²</li> <li>3. <b>Labuan</b> – 92 km²</li> </ol> <p>Walaupun Labuan lebih besar daripada Putrajaya dari segi keluasan, Putrajaya dianggap sebagai Wilayah Persekutuan kedua terbesar dari segi kepentingan pentadbiran dan pembangunan. Kuala Lumpur kekal sebagai yang terbesar dan paling maju.</p> <p>✗</p>	<p>Wilayah Persekutuan yang kedua terbesar di Malaysia</p> <p>Berdasarkan keluasan, Wilayah Persekutuan yang kedua terbesar di Malaysia ialah <b>Labuan</b>.</p> <p>Berikut adalah kedudukan Wilayah Persekutuan mengikut saiz:</p> <ol style="list-style-type: none"> <li>1. <b>Kuala Lumpur</b> (243.65 km²)</li> <li>2. <b>Labuan</b> (98 km²)</li> <li>3. <b>Putrajaya</b> (49 km²)</li> </ol> <p>✓</p>	<p>Wilayah Persekutuan yang kedua terbesar di Malaysia</p> <p><b>ILMUchat</b> Wilayah Persekutuan yang kedua terbesar di Malaysia ialah <b>Wilayah Persekutuan Labuan</b>.</p> <p>Labuan, yang terletak di luar pantai Sabah, meliputi kawasan seluas 91.64 kilometer persegi. Ia diisytiharkan sebagai Wilayah Persekutuan kedua pada 16 April 1984, selepas Kuala Lumpur. Walaupun lebih kecil berbanding Kuala Lumpur dari segi keluasan, Labuan memainkan peranan penting sebagai pusat kewangan antarabangsa dan hab logistik, serta terkenal dengan industri minyak dan gas.</p> <p>✓</p>
(a) DeepSeek	(b) Gemini 2.5 Flash	(c) ILMUchat

Figure 11: Comparison of responses to factual queries about the Malaysian context, highlighting differences between DeepSeek, Gemini 2.5 Flash and ILMUchat.



## 2.3 Vision Capability

ILMUchat brings multimodal capabilities that allow it to interpret and respond to visual inputs in a way that is culturally grounded in the Malaysian context. By integrating image understanding into its language model, ILMUchat can recognise and describe visual content related to Malaysian food, culture, and landmarks, reflecting not just technical proficiency, but deep alignment with the everyday visual environment familiar to Malaysians.

**ILMUchat understands Malaysian context through visual inputs.** As a multiracial and culturally diverse nation, Malaysia presents unique visual cues that span cuisine, traditional practices, and architectural landmarks. ILMUchat is designed to recognise and interpret these cues accurately. For instance, as shown in Figure 12, ILMUchat can identify iconic local dishes such as *Nasi Lemak* and *Char Kway Teow*. It also understands traditional Malaysian games like *congkak* and *wau*, as illustrated in Figure 13, and correctly recognises national landmarks such as Batu Caves and the Petronas Twin Towers (Figure 14).

**ILMUchat also delivers culturally rich visual descriptions that match the quality of leading models like ChatGPT.** As demonstrated in Figure 15, ILMUchat generates detailed and culturally grounded responses to visual prompts. For example, both ILMUchat and ChatGPT correctly identify that the green colour of *ondeh-ondeh* comes from pandan (screwpine) leaves, while Gemini omits this culturally specific detail. These results highlight ILMUchat’s strong performance not only in object recognition but also in the contextual understanding needed to make its responses relevant and authentic to Malaysian users.

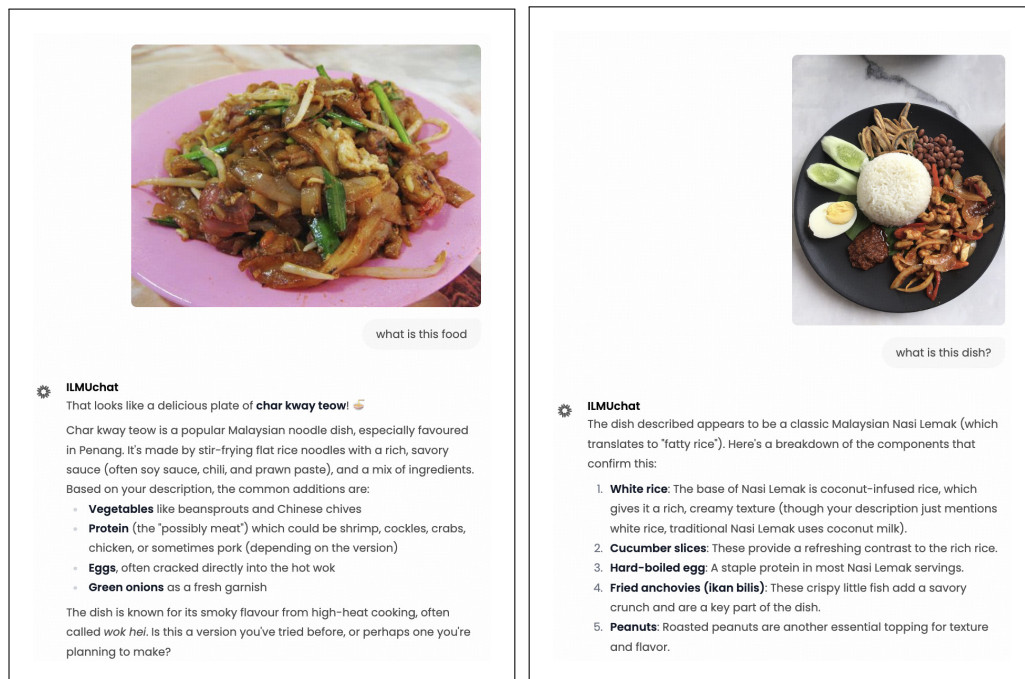
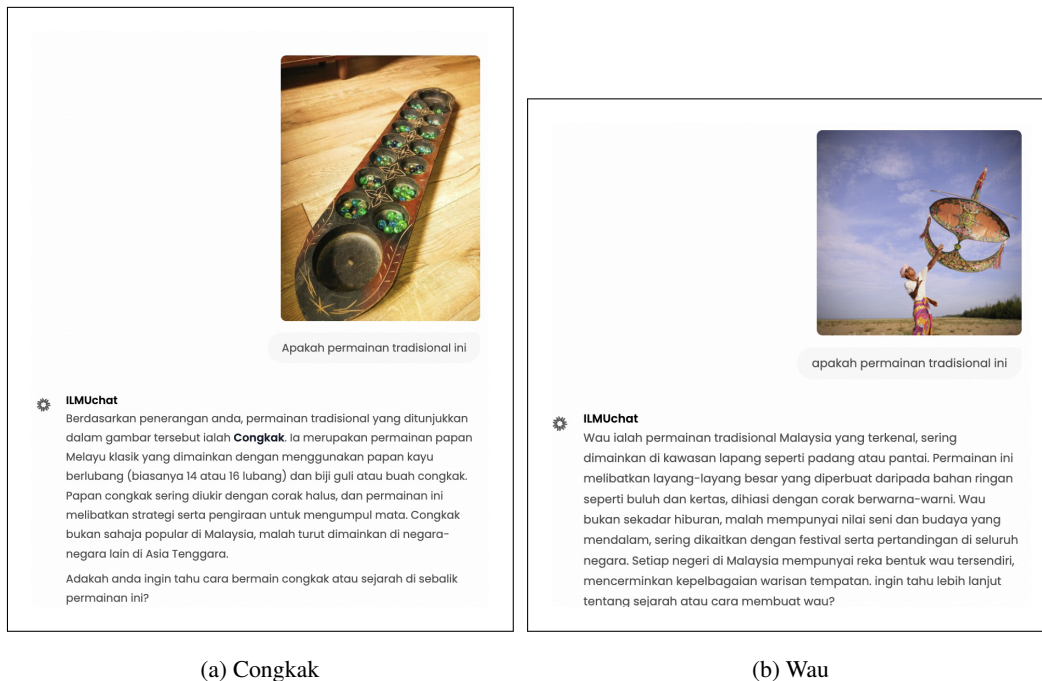


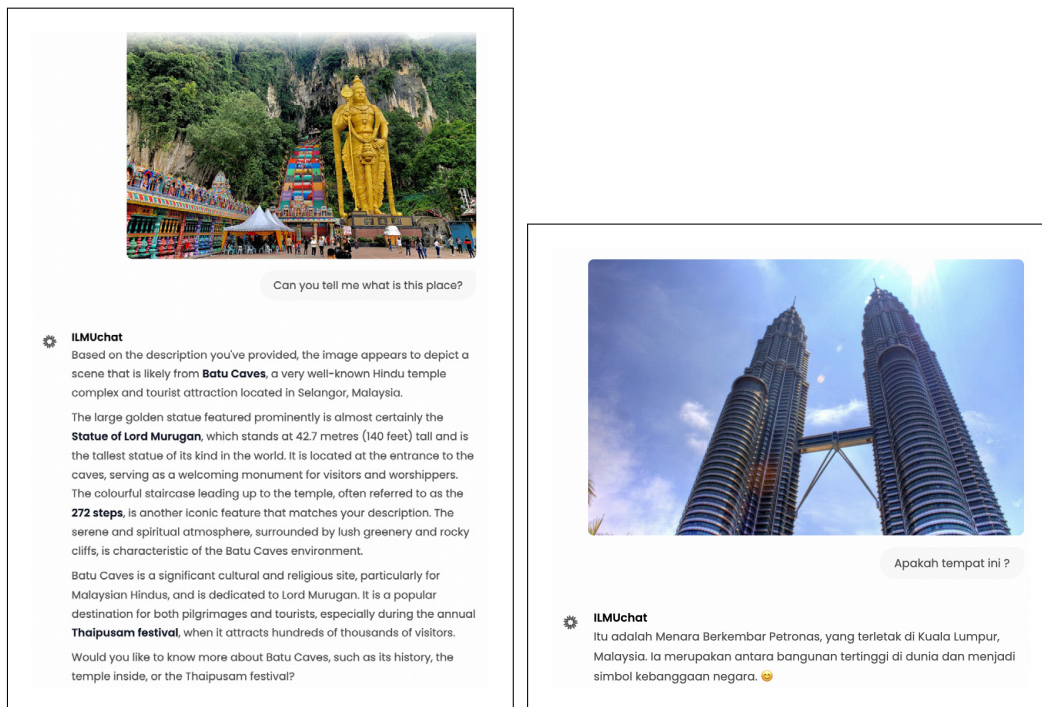
Figure 12: Local Food: Through its visual understanding module, ILMUchat is able to accurately recognize culturally significant Malaysian foods, including *Char Kway Teow* and *Nasi Lemak*.



(a) Congkak

(b) Wau




Figure 13: Traditional Games: ILMUchat demonstrates visual recognition of traditional Malaysian games, including (a) *congkak* and (b) *wau*, reflecting cultural awareness in image understanding.






(a) Batu Caves

(b) Petronas Twin Towers

Figure 14: Landmark: ILMUchat can accurately recognize iconic Malaysian landmarks such as (a) Batu Caves and the (b) Petronas Twin Towers through visual input.

 <p><b>Analysis</b></p> <p>This is a vibrant, textured painting of "Satay," a popular Southeast Asian dish. The artwork depicts four skewers of grilled meat, likely chicken or beef, charred and caramelized.</p> <p>To the side, there is a bowl of rich, creamy peanut sauce, which is a key accompaniment to satay. The dish is served with traditional sides, including sliced cucumbers, cubes of what appear to be rice cakes (ketupat), and red onions. The overall style is expressive and painterly, with visible brushstrokes that give the image a lot of character and a warm, inviting feel. The word "SATAY" is painted in large, yellow letters at the top, making the subject of the painting clear.</p>	 <p>This is a vibrant painting of Satay, a popular Southeast Asian dish. The artwork captures:</p> <ul style="list-style-type: none"> <li>• Grilled skewers of marinated meat (likely chicken or beef), charred and caramelized.</li> <li>• Peanut sauce in a bowl, thick and rich, typically served as a dip.</li> <li>• Traditional sides: <ul style="list-style-type: none"> <li>• Cubed rice cakes (ketupat or nasi impit)</li> <li>• Slices of cucumber</li> <li>• Chunks of red onion</li> </ul> </li> <li>• All of this is served on a banana leaf over a round plate, giving it an authentic presentation.</li> </ul> <p>The warm, textured brushstrokes and bold use of colors (orange, green, and brown) evoke a homely, rustic feel—emphasizing the dish's comforting and flavorful nature.</p>	 <p><b>ILMUchat</b></p> <p><b>Satay: A Malaysian Delight!</b></p> <p>The image beautifully captures a classic Malaysian satay dish, showcasing skewers of perfectly grilled meat served alongside traditional accompaniments. This beloved dish is a staple in Malaysian cuisine, enjoyed by people of all backgrounds.</p> <p>Satay typically consists of marinated meat (often chicken, beef, or mutton) skewered and grilled over hot coals, which gives it that distinctive smoky flavour. The dish is usually served with a creamy peanut dipping sauce, known as kachang, that balances the richness of the grilled meat with its tangy and slightly sweet notes.</p> <p>The accompaniments in your image – crisp cucumber slices and cubes of tofu – provide a refreshing contrast to the rich, savoury satay. These are common additions that help cleanse the palate between bites.</p> <p>Satay is more than just food; it's a cultural experience often enjoyed at celebrations, night markets, and family gatherings. The word 'satay' itself actually comes from the Minangkabau language in Indonesia, reflecting the dish's Southeast Asian roots.</p> <p>Would you like to know more about the regional variations of satay in Malaysia or how to make the perfect satay marinade at home?</p>
Gemini	ChatGPT	ILMUchat

(a) Satay

 <p><b>Analysis</b></p> <p>Based on the image provided, the painting depicts a plate of "ondeh-ondeh." These are small, round, green-colored glutinous rice balls, covered in what appears to be grated coconut. The central one has been cut open, revealing a melted, brown sugar filling. The style of the painting is textured and impressionistic. The title "ONDEH-ONDEH" is prominently displayed at the top of the image.</p>	 <p>The image you've shared is a stylized painting of <b>ondeh-ondeh</b>, a traditional Southeast Asian dessert popular in Malaysia and Indonesia.</p> <p>Here's a brief description of <b>ondeh-ondeh</b>:</p> <p><b>Onde-Onde</b> (also spelled <b>onde-onde</b> or <b>klepon</b> in Indonesia) are chewy, glutinous rice balls typically made from <b>pandan-flavored glutinous rice flour</b>, giving them their distinct green color. They are coated in shredded coconut and <b>filled with melted palm sugar (gula Melaka)</b>, which bursts in your mouth when bitten into.</p>	 <p><b>ILMUchat</b></p> <p>Thank you for sharing the image description! It sounds like a delightful and detailed painting capturing the traditional sweet dumpling known as ondeh-ondeh. The fact that one is depicted cut open to show the filling really highlights the artist's attention to detail.</p> <p>Onde-ondeh are indeed a popular treat, often enjoyed during various celebrations in Malaysia and Indonesia. They typically feature a chewy outer layer made from glutinous rice flour and a palm sugar filling, which melts when cooked, creating a sweet surprise inside. The green colour usually comes from pandan juice, giving them a distinctive aroma.</p> <p>Is there anything specific about the painting or the ondeh-ondeh itself that you'd like to discuss further?</p>
Gemini	ChatGPT	ILMUchat

(b) Ondeh-ondoh

Figure 15: Comparison of vision-language responses from Gemini, ChatGPT, and ILMUchat, each shown images of *satay* and *ondeh-ondoh* to assess cultural grounding and descriptive accuracy.

## 2.4 Audio Capabilities

Spoken communication in Malaysia is often characterized by informal phrasing, diverse accents, and code-switching, where speakers fluidly alternate between Malay and English within a single utterance. These features are common in daily interactions but pose challenges for conventional audio-based systems, which are typically trained on monolingual data and standardized accents.

ILMUchat is designed to handle this linguistic complexity. It demonstrates strong performance in transcribing and producing natural Malaysian speech, including accented delivery, informal expressions, and mixed-language usage. Unlike standard audio-based systems, ILMUchat accurately captures the structure and vocabulary of Malaysian code-switched utterances without distortion or language bias.

Figures 16 and 17 provide side-by-side comparisons of transcriptions generated by ILMUchat, GPT-4o-mini-transcribe, and GPT-4o-transcribe for a sample utterance containing both English and Malay. ILMUchat consistently preserves the full meaning and form of both languages, reflecting a more contextually faithful transcription of multilingual speech.

This kind of code-switching is just one aspect of the broader variability found in Malaysian spoken language. **Significant accentual and phonological variation** also exists within Malay itself, shaped by the country’s multi-ethnic and regionally diverse population. Among native Malay speakers, regional influences such as those from *Kelantanese*, *Terengganu*, and *Johorean* Malay result in distinct differences in vowel quality, prosody, intonation, and phonotactic patterns. In parallel, non-Malay ethnic communities, including Chinese-Malaysians and Indian-Malaysians, exhibit systematic phonetic transfer when speaking Malay as a second language. For example, *Chinese-Malaysian* speakers may reduce vowel length contrasts, simplify final consonants, and carry over tonal influences from Mandarin. *Indian-Malaysian* speakers, on the other hand, often introduce retroflex consonants, intervocalic voicing, and stress-timed rhythm patterns influenced by Tamil.

Adding further complexity, informal Malay, particularly *Bahasa Pasar* introduces widespread elision and assimilation phenomena, affecting both native and non-native speakers. These combined effects of regional, ethnic, and stylistic variation create significant challenges for conventional audio systems trained primarily on standardized Malay speech.

Despite this diversity, Figures 18, 19 and 20 show that ILMUchat demonstrates strong robustness across these variants, maintaining high transcription accuracy even in the presence of pronounced dialectal and phonological variation.

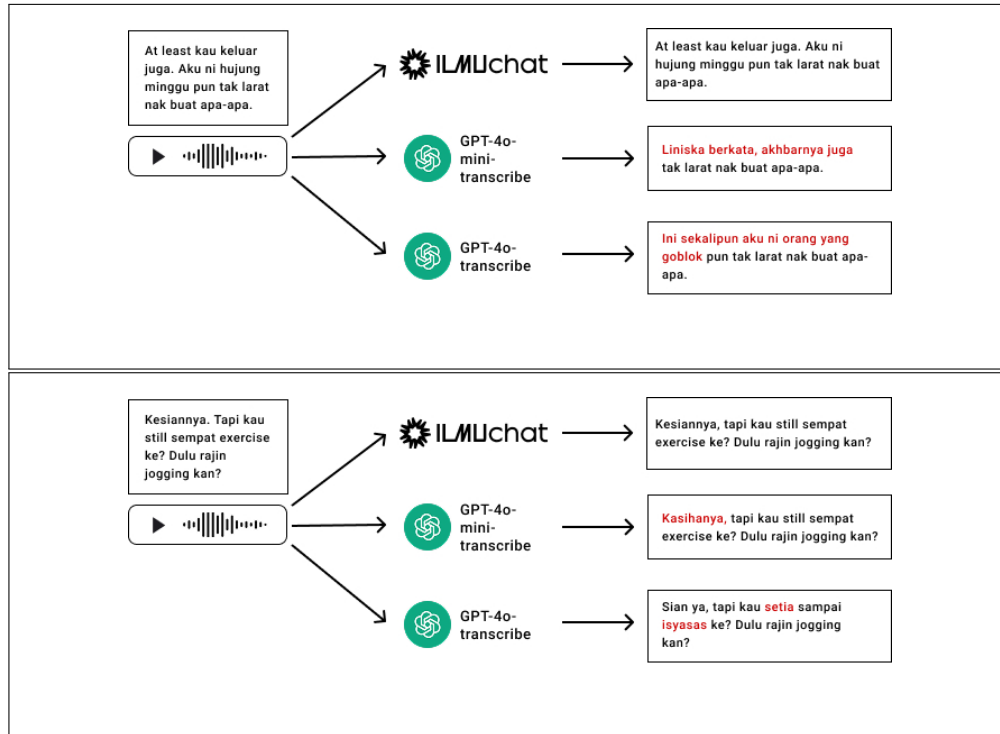


Figure 16: Code-Switching: Comparison of ASR Outputs from GPT Models and ILMUchat for a code-switched utterance by a *female* speaker

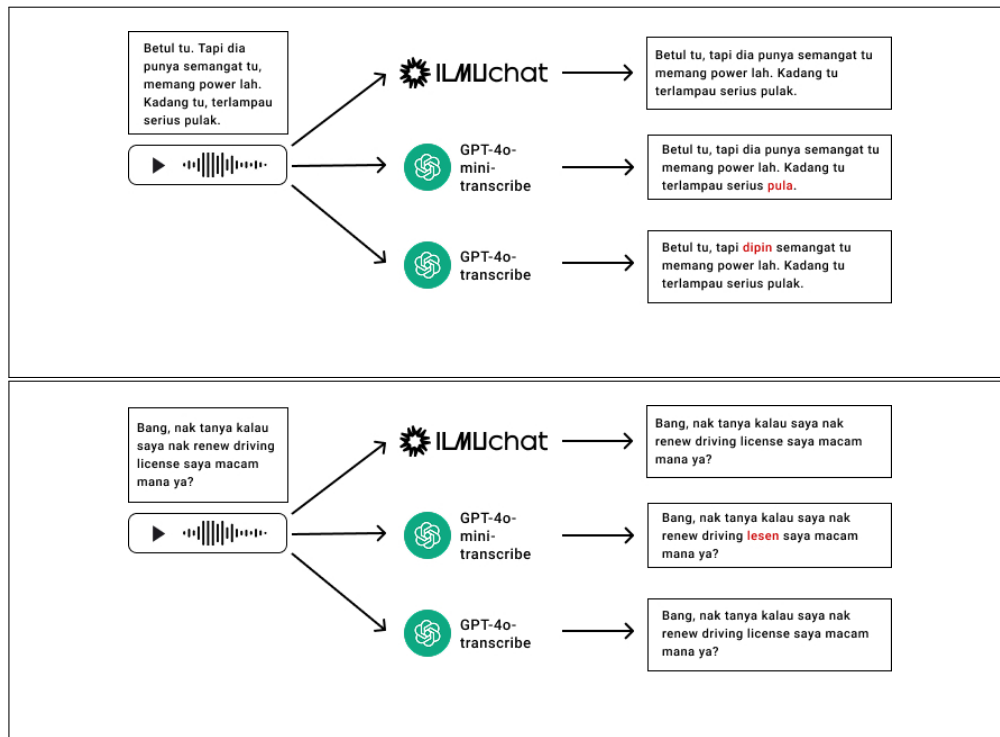


Figure 17: Code-Switching: Comparison of ASR Outputs from GPT Models and ILMUchat for a code-switched utterance by a *male* speaker

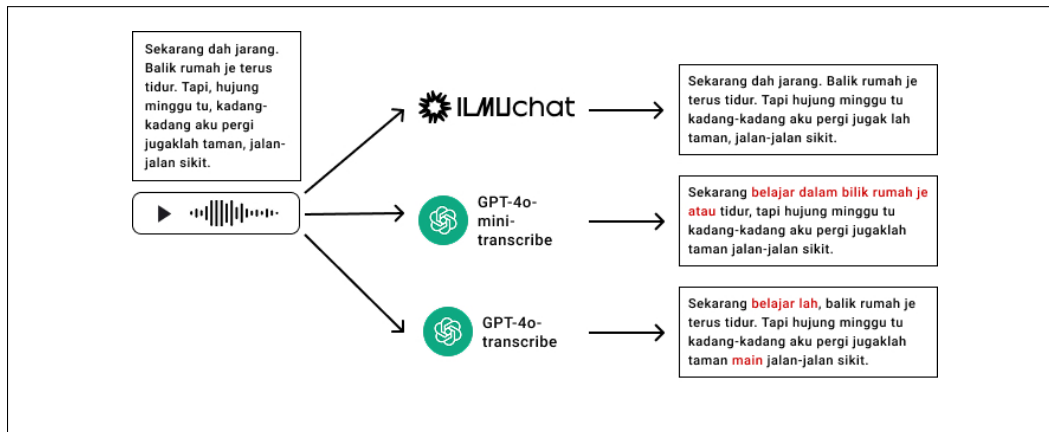


Figure 18: Malay Ethnic Speaker: ASR performance comparison between ILMUchat and GPT models for a *Malay*-spoken utterance.

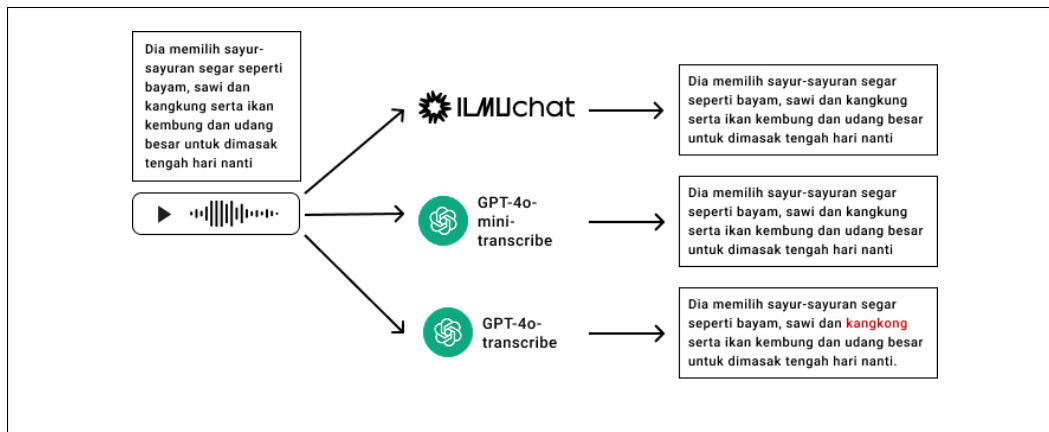


Figure 19: Indian Ethnic Speaker: ASR performance comparison between ILMUchat and GPT models for a *Indian*-spoken utterance.

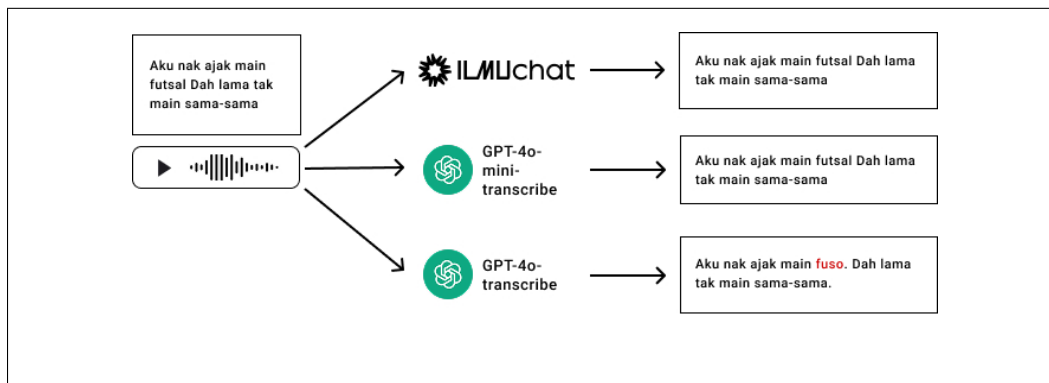


Figure 20: Chinese Ethnic Speaker: ASR performance comparison between ILMUchat and GPT models for a *Chinese*-spoken utterance.



### 3 Quantitative Evaluations

To assess ILMU’s general language understanding and cultural alignment, we evaluate its performance on several standard multitask benchmarks. As shown in Table 2, ILMU exhibits strong and competitive performance across three prominent benchmarks: MMLU [2], CMMLU [3], and MalayMMLU [5]. Notably, ILMU achieves the highest score on MalayMMLU, surpassing leading models such as GPT-5/4o, LLaMA-4-Scout, GLM-4.5 and DeepSeek-v3.

MalayMMLU [5] is the *first* dedicated multitask language-understanding benchmark for Bahasa Melayu, comprising 24,213 questions spanning Malaysia’s primary (Years 1–6) and secondary (Forms 1–5) curricula. It covers five broad domains split across 22 subjects, providing a standardized, curriculum-aligned evaluation of LLMs in Malay language and Malaysian cultural contexts.

As detailed in Table 3, on MalayMMLU’s zero-shot, first-token accuracy evaluation, ILMU achieves state-of-the-art performance with a macro average of 87.20%, leading every domain: Language (89.36%), Humanities (88.40%), STEM (88.05%), Social Science (85.44%), and Others (85.08%). The narrow cross-domain spread (~4.3%) reflects strong cross-task generalization rather than subject-specific overfitting. This benchmark evaluation covered 54 models from 20 leading LLM developers, including the latest GPT-5<sup>2</sup>, LLaMA-4-Scout<sup>3</sup>, and GLM-4.5<sup>4</sup>, representing both closed- and open-source paradigms. Compared to the next-best closed model (GPT-4o, 84.97% average), ILMU delivers a +2.23% gain and maintains double-digit margins over most open-source baselines (e.g., LLaMA, Qwen, Mistral). On the Bahasa Melayu-first, Malaysia-grounded benchmark, ILMU is therefore the clear SOTA, combining high macro accuracy, domain consistency, and robust Malay-centric reasoning. In other words, this performance highlights ILMU’s deep linguistic proficiency in the Malay language and its close alignment with Malaysian culture, reinforcing its suitability for local applications.

Nonetheless, ILMU is competitive with top-tier models on both MMLU [2] and CMMLU [3] benchmarks, occasionally edging them out. Its solid performance in Malay, Chinese, and English positions it as a uniquely capable assistant for Malaysia’s multilingual and multicultural landscape.

### 4 Conclusion

ILMUchat represents a major advancement in Malaysia’s AI landscape. It is a multimodal generative assistant developed entirely within the country, designed to serve the linguistic and cultural needs of its diverse population. Built upon the ILMU foundation model, trained from scratch using Malaysia-centric data, ILMUchat offers locally grounded, contextually intelligent responses across text, vision, and speech modalities.

Evaluations across leading benchmarks, including MMLU, CMMLU and MalayMMLU demonstrate ILMU’s strong multilingual capabilities. The model achieves state-of-the-art performance on MalayMMLU, reflecting its deep alignment with national language standards.

Beyond benchmark performance, ILMUchat demonstrates robust real-world functionality, from accurately transcribing informal and code-switched speech to recognizing local foods, traditional games, and landmarks. It also adapts effectively to diverse styles of Malay, including literary, colloquial, and royal variants, affirming its cultural sensitivity and linguistic versatility.

Support for informal and regional language use is an area of active development. While comprehensive dialect coverage is still being extended, ongoing updates aim to enhance ILMUchat’s ability to reflect Malaysia’s full linguistic and cultural diversity. These improvements are guided by community feedback and a continuous learning pipeline designed to ensure long-term relevance and inclusivity.

Looking ahead, ILMUchat affirms the importance of building AI technologies that reflect Malaysia’s unique cultural, linguistic, and societal fabric. We envision this work as a catalyst for national progress and invite collaboration across local universities, industry partners, government agencies, and civil society. By working together, we can shape an AI future that is not only intelligent and capable but proudly Malaysian in identity, values, and purpose. *Malaysia Boleh.*

<sup>2</sup><https://openai.com/index/introducing-gpt-5/>

<sup>3</sup><https://huggingface.co/meta-llama/Llama-4-Scout-17B-16E>

<sup>4</sup><https://docs.z.ai/guides/llm/glm-4.5>

Table 2: Comparison of ILMU and other models on Multitask Language Understanding (MLU) benchmarks, including MMLU, CMMLU, and MalayMMLU. **Bold** indicates the best score.

Model	MMLU	CMMLU	MalayMMLU
GPT-4o	83.94	76.53	84.97
DeepSeek-V3	<b>84.15</b>	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
ILMU (ours)	80.39	<b>83.64</b>	<b>87.20</b>

Table 3: Zero-shot results of various LLMs/LVLMs on MalayMMLU based on *first token accuracies*. Highest scores are **bolded**. <sup>†</sup> denotes LLMs finetuned with SEA datasets. <sup>††</sup> denotes open-source GLM-4 which is different from the closed-source GLM-4.

Organization	Model	Language	Humanities	STEM	Social Science	Others	Average
		Acc.	Acc.	Acc.	Acc.	Acc.	Acc.
	Random	38.01	42.09	36.31	36.01	38.07	38.02
YTL AI Labs	ILMU	<b>89.36</b>	<b>88.40</b>	<b>88.05</b>	<b>85.44</b>	<b>85.08</b>	<b>87.20</b>
Meta	LLaMA-4-Scout (17B-16E)	78.69	78.27	77.98	73.55	75.58	76.54
	LLaMA-3.2 (3B)	58.52	60.66	56.65	54.06	52.75	56.45
	LLaMA-3.2 (1B)	38.88	43.30	40.65	40.56	39.55	40.46
	LLaMA-3.1 (70B)	78.75	82.59	78.96	77.20	75.32	78.44
	LLaMA-3.1 (8B)	65.47	67.17	64.10	62.59	62.13	64.24
	LLaMA-3 (8B)	63.93	66.21	62.26	62.97	61.38	63.46
	LLaMA-2 (13B)	45.58	50.72	44.13	44.55	40.87	45.26
Qwen (Alibaba)	LLaMA-2 (7B)	47.47	52.74	48.71	50.72	48.19	49.61
	Qwen-2.5 (72B)	79.09	79.95	80.88	75.80	75.05	77.79
	Qwen-2.5 (32B)	76.96	76.70	79.74	72.35	70.88	74.83
	Qwen-2-VL (7B)	68.16	63.62	67.58	60.38	59.08	63.49
	Qwen-2-VL (2B)	58.22	55.56	57.51	53.67	55.10	55.83
	Qwen-1.5 (14B)	64.47	60.64	61.97	57.66	58.05	60.47
	Qwen-1.5 (7B)	60.13	59.14	58.62	54.26	54.67	57.18
Zhipu	Qwen-1.5 (4B)	48.39	52.01	51.37	50.00	49.10	49.93
	Qwen-1.5 (1.8B)	42.70	43.37	43.68	43.12	44.42	43.34
	GLM-4.5 (Thinking)	84.32	86.14	84.40	81.01	82.27	83.36
	GLM-4.5 (Non-Thinking)	81.23	81.50	83.71	78.38	76.97	79.98
	GLM-4-Plus	78.04	75.63	77.49	74.07	72.66	75.48
	GLM-4-Air	67.88	69.56	70.20	66.06	66.18	67.60
	GLM-4-Flash	63.52	65.69	66.31	63.21	63.59	64.12
OpenAI	GLM-4	63.39	56.72	54.40	57.24	55.00	58.07
	GLM-4 <sup>††</sup> (9B)	58.51	60.48	56.32	55.04	53.97	56.87
	GPT-5	83.59	76.50	78.10	80.73	75.44	79.53
	GPT-4o	87.64	87.78	83.54	82.84	82.34	84.97
	GPT-4	82.90	83.91	78.80	77.29	77.33	80.11
Mistral	GPT-4o mini	82.03	81.50	78.51	75.67	76.30	78.78
	GPT-3.5	69.62	71.01	67.17	66.70	63.73	67.78
	Pixtral (12B)	64.81	62.68	64.72	63.93	59.49	63.25
	Mistral Small (22B)	65.19	65.03	63.36	61.58	59.99	63.05
Google	Mistral-v0.3 (7B)	56.97	59.29	57.14	58.28	56.56	57.71
	Mistral-v0.2 (7B)	56.23	59.86	57.10	56.65	55.22	56.92
	Gemma-2 (9B)	75.83	72.83	75.07	69.72	70.33	72.51
Mesolitica	Gemma (7B)	45.53	50.92	46.13	47.33	46.27	47.21
	Gemma (2B)	46.50	51.15	49.20	48.06	48.79	48.46
	MaLLaM-v2.5 small	73.00	71.00	70.00	72.00	70.00	71.53
Microsoft	MaLLaM-v2 <sup>†</sup> (5B)	42.57	46.44	42.24	40.82	38.74	42.08
	Phi-3 (14B)	60.07	58.89	60.91	58.73	55.24	58.72
SAIL (Sea)	Phi-3 (3.8B)	52.24	55.52	54.81	53.70	51.74	53.43
	Sailor <sup>†</sup> (14B)	78.40	72.88	69.63	69.47	68.67	72.29
Stability AI	Sailor <sup>†</sup> (7B)	74.54	68.62	62.79	64.69	63.61	67.58
	StableLM 2 (12B)	53.40	54.84	51.45	51.79	50.16	52.45
01.AI	StableLM 2 (1.6B)	43.92	51.10	45.27	46.14	46.75	46.48
	Yi-1.5 (9B)	56.20	53.36	57.47	50.53	49.75	53.08
Agmo	Merdeka-LLM	56.92	60.36	57.63	56.82	55.10	57.28
AI Singapore	SEA-LION v3 (70B)	79.20	81.11	78.76	76.45	75.22	78.03
Baichuan	Baichuan-2 (7B)	40.41	47.35	44.37	46.33	43.54	44.30
Cohere for AI	Command R (32B)	71.68	71.49	66.68	67.19	63.64	68.47
Damo (Alibaba)	SeaLLM-v2.5 <sup>†</sup> (7B)	69.75	67.94	65.29	62.66	63.61	65.89
Deepseek	DeepSeek-V3	83.13	78.84	83.91	78.25	78.00	80.56
GoTo + IndoSat	Sahabat-AI-v2 (70B)	80.60	81.25	78.31	76.39	74.93	78.31
OpenGVLab	InternVL2 (40B)	70.36	68.49	64.88	65.93	60.54	66.51
Technology Innovation Institute (TII)	Falcon3 (10B)	54.77	60.17	58.20	56.76	54.04	56.38
Yellow.ai	Komodo <sup>†</sup> (7B)	43.62	45.53	39.34	39.75	39.48	41.72

## Limitations

Like all foundation models, ILMUchat may occasionally generate responses that are factually inaccurate despite appearing plausible, a phenomenon known as *hallucination*. It may also face challenges with causal reasoning, complex logical deductions, and counterfactual understanding. While proficient in multiple languages, its strongest capabilities are specifically in the Malay language. Ongoing development efforts continue to address these limitations.

## References

- [1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.
- [2] Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*, 2020.
- [3] Haonan Li, Yixuan Zhang, Fajri Koto, Yifei Yang, Hai Zhao, Yeyun Gong, Nan Duan, and Timothy Baldwin. Cmmlu: Measuring massive multitask language understanding in chinese. *arXiv preprint arXiv:2306.09212*, 2023.
- [4] David Ong and Peerat Limkonchotiawat. Sea-lion (southeast asian languages in one network): a family of southeast asian language models. In *Proceedings of the 3rd workshop for natural language processing open source software (NLP-OSS 2023)*, pages 245–245, 2023.
- [5] Soon Chang Poh and et al. Malaymmlu: A multitask benchmark for the low-resource malay language. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 650–669, 2024.
- [6] Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, et al. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.
- [7] An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025.
- [8] Husein Zolkepli, Aisyah Razak, Kamarul Adha, and Ariff Nazhan. Mallam–malaysia large language model. *arXiv preprint arXiv:2401.14680*, 2024.