

Digital Transformation in ASEAN Humanities: Integration of AI and Language Technologies for Multilingual Information Retrieval

Qi Zhu

*Qi Zhu, International College for Chinese Studies,
Nanjing Normal University, Nanjing, Jiangsu, China
absinthzhu@163.com*

and

Zijun Shen*

*Department of Foreign Languages, Sichuan University of
Media and Communications, Chengdu, Sichuan, China
<https://orcid.org/0009-0008-4170-6524>
prof.shen_zijun@whu.edu.cn*

Abstract

The digital transformation of the humanities within ASEAN has been influenced by artificial intelligence (AI) and language technologies, particularly in the domain of multilingual information retrieval. This research aimed to examine the impact of AI and language technologies on multilingual information retrieval across ASEAN nations. Data were gathered from academic staff at four universities in Malaysia, Indonesia, Thailand, and Brunei, with a total sample of 357 respondents. The data were analysed using SPSS 29 and Smart PLS 4. The findings substantiated the significant influence of AI and language technologies on multilingual information retrieval within ASEAN countries. It is recommended that ASEAN nations prioritise digital transformation initiatives, particularly the integration of AI and language technologies in multilingual information retrieval. Future advancements should ensure that AI fosters inclusive and multilingual access to knowledge while preserving ASEAN's linguistic diversity.

Keywords: Artificial Intelligence, Linguistic Diversity, Digital Humanities, Multilingual Information Retrieval.

Introduction

The past few years have witnessed significant advancements in artificial intelligence (AI) and language technologies, which are increasingly being integrated into the digital transformation of the humanities across various regions, including ASEAN (Association of Southeast Asian Nations). Mustapa et al. (2023) highlight a departure from traditional resource-sharing methods, particularly in the preservation, accessibility, and dissemination of cultural and educational resources in ASEAN, marking a new phase in digital evolution. The region's linguistic diversity, encompassing distinct languages, dialects, and socio-cultural contexts, presents both opportunities and challenges for digital technologies, particularly in multilingual information retrieval (Nzomo et al., 2021). Such retrieval systems are essential in ensuring equitable access to cultural and educational content while mitigating linguistic barriers (Alon and Krtalić, 2025). Given AI's advanced capabilities in language processing, its application in multilingual information retrieval is highly beneficial. However, challenges such as multilingualism, context sensitivity, and data accessibility must be addressed to optimise information access in ASEAN's humanities, leveraging AI's potential.

AI-powered language technologies have become indispensable tools for multilingual information retrieval in cultural and educational resources (Horna-Saldaña et al., 2025). The primary languages in ASEAN include Thai, Malay, Vietnamese, and Tagalog, among others, and AI technologies can facilitate cultural exchange and mutual understanding across the region. Consequently, users from diverse linguistic backgrounds can access and

engage with cultural and academic content more effectively (Heng et al., 2022). By integrating AI into information retrieval systems, scholars and students gain access to previously inaccessible information, overcoming linguistic and geographical barriers. However, for these technologies to be effectively implemented, they must be tailored to the region's linguistic and cultural diversity (Lee et al., 2023). As Alsayed et al. (2023) argue, dialectal variations and syntactical differences often pose significant challenges to language models, necessitating the development of regionally adapted multilingual datasets.

The advancement of AI-driven language technologies in ASEAN's humanities also raises concerns regarding cultural preservation and linguistic diversity. Machine learning algorithms frequently struggle to interpret culturally nuanced expressions, idiomatic phrases, and the deeper meanings embedded within regional languages (Khan et al., 2023). This issue extends to the academic taxonomy within conceptual domains, which must be preserved within digital systems to safeguard intangible cultural heritage. Given that the region encompasses over 800 indigenous languages, scholars emphasise the necessity of cultural sensitivity in AI development to prevent the erosion of linguistic diversity. By aligning AI technology with cultural considerations, ASEAN countries can leverage digital advancements to preserve their historical and cultural heritage.

Furthermore, AI and language technologies in education hold the potential to democratise knowledge access, particularly for learners in remote or underserved ASEAN areas, by providing access to global resources (Horna-Saldana et al., 2025). AI-driven digital platforms now offer cross-linguistic research tools that enable students to retrieve information across multiple languages with greater ease (Safdar et al., 2020). Additionally, AI is increasingly employed in content curation and recommendation systems to deliver personalised learning experiences, thereby enhancing the accessibility of educational content for students from diverse backgrounds (An et al., 2023). However, to fully realise these benefits, collaborative efforts among governments, educational institutions, and technology developers are required to prevent further exacerbation of the digital divide. It is essential to promote educational inclusivity by ensuring equitable access to AI-driven educational resources, fostering an environment where students can thrive within the globally interconnected academic landscape (Han et al., 2024).

Although prior research has explored AI's effectiveness in information retrieval from digital platforms, its role in multilingual environments, particularly within ASEAN, remains underexamined. Consequently, there is limited scholarly discourse on the relationship between AI, language

technologies, and multilingual information retrieval. This research aimed to investigate the impact of AI and language technologies on multilingual information retrieval within ASEAN nations, addressing a critical gap in the literature while offering practical insights into AI integration for overcoming language learning barriers. Data were collected from academic staff at four universities in Malaysia, Indonesia, Thailand, and Brunei, yielding a sample of 357 respondents. The data were analysed using SPSS 29 and Smart PLS 4. The findings confirmed the significant impact of AI and language technologies on multilingual information retrieval in ASEAN countries. It is recommended that ASEAN nations prioritise digital transformation initiatives, particularly the integration of AI and language technologies in multilingual information retrieval. This study is structured around a comprehensive literature review, a detailed methodological framework, data analysis, discussion of findings, and practical implications. Additionally, study limitations and future research directions are outlined.

Literature Review

Theoretical Foundation

The Technology Acceptance Model (TAM) serves as a valuable framework for analysing the adoption and utilisation of AI and language technologies in multilingual information retrieval. The model's core constructs—perceived usefulness (PU) and perceived ease of use (PEU)—explain how various user groups, including researchers, businesses, and multilingual communities, evaluate and accept AI-driven language technologies for multilingual information retrieval (Liu and Ma, 2024). When AI-powered language tools are perceived as enhancing search accuracy and efficiency across multiple languages (PU) and are user-friendly with minimal learning barriers (PEU), users are more inclined to adopt them (Lin and Yu, 2025). Moreover, external factors such as the availability of training, prior experience with AI technologies, and system reliability may influence both behavioural intention and actual system usage. By integrating TAM, this study offers insights into the acceptance and adoption of AI-driven language technologies, contributing to the development of more effective and user-centric multilingual information retrieval systems (Chonka et al., 2023).

Artificial Intelligence and Multilingual Information Retrieval

In contemporary times, AI-based technologies have become instrumental in information retrieval and document

access (Alothman and Wahab Sait, 2022). These technologies facilitate academics and students in acquiring knowledge by extracting relevant information from diverse sources. AI possesses the capability to comprehend user queries and implement retrieval processes tailored to the specified query parameters (Yang et al., 2022). The evolution of artificial intelligence is enhanced through continuous training, enabling the accurate retrieval of required information (Siciliani et al., 2024). AI plays a crucial role in information management systems, where its core function in retrieval aids in interpreting user queries and delivering precise information (Huang and Huang, 2024).

With advancements in AI technology, information retrieval has become more efficient, leveraging language models to process linguistic variations across different cultures and contexts (Safdar et al., 2020). However, a comprehensive understanding of AI-based systems is essential to enhance the accuracy of information delivery and support effective resource allocation (Ayotunde et al., 2023). Furthermore, AI contributes to information management by categorising practical knowledge into distinct sections based on its learned models, ensuring that accurate and relevant information is directed to the appropriate users. AI developers must integrate cultural and linguistic considerations to refine interpretation and query processing across multiple languages (Lee and Kim, 2023). This approach advances AI development by ensuring the retrieval of information from diverse sources while maintaining transparency and proper information management (de Oliveira et al., 2023). Moreover, AI tools must uphold the responsibility of delivering precise and relevant information in real time (Fan et al., 2022). Consequently, advanced training of AI systems is imperative to refine retrieval processes and categorise information according to subject areas, thereby promoting transparency in shared knowledge (Zhu et al., 2024). The integration of AI in information management has become increasingly significant, particularly in handling multilingual content (Ahmad et al., 2023). To enhance AI productivity, developers must optimise its efficiency in processing diverse languages and cultural contexts (Han et al., 2024). Such advancements enable AI tools to interpret queries accurately and generate precise responses aligned with users' informational needs.

H1: There is a relationship between artificial intelligence and multilingual information retrieval.

Language Technologies and Multilingual Information Retrieval

Similar to artificial intelligence, language technologies, particularly machine learning-driven

intelligence mechanisms, play a crucial role in information retrieval (Alsalmi, 2021). These technologies facilitate the categorisation and interpretation of information according to user requirements (Sansone and Sperlí, 2022). For AI development, enhancing language comprehension within machines is essential to ensure accurate responses aligned with user demands. The increasing integration of technology into modern life has led to widespread adoption of language-based technologies and models in workplaces for information retrieval (Kim et al., 2024). However, AI development necessitates a machine learning approach that processes data across various languages while considering cultural values (Ali and Farha, 2025).

Moreover, machine learning is instrumental in delivering high-quality information, a critical requirement for AI systems (Efimov et al., 2023). While information retrieval is inherently complex, machine learning models simplify the process by retrieving data based on predefined queries (Barsha and Munshi, 2024). AI-driven machine learning mechanisms are particularly vital for information sharing, fostering improved opportunities for data retrieval (Khurana et al., 2023). However, effective utilisation of AI-powered information retrieval requires users to possess a foundational understanding of query structuring and database navigation (Khan et al., 2024). The integration of machine learning and information management systems enhances information sharing and retrieval processes (Jantscher et al., 2023). Training language models to optimise information retrieval across multiple languages can enhance the productivity of both businesses and individuals (Dash, 2024). Within academic institutions, language technologies can facilitate seamless knowledge retrieval from diverse sources and languages, mitigating cultural and technical barriers (Litschko et al., 2022). Consequently, AI-based programming for information retrieval is essential in advancing comprehension and development within multilingual environments (Lawrie et al., 2023). Furthermore, AI-driven information retrieval systems significantly contribute to the efficiency and accuracy of knowledge dissemination.

H2: There is a relationship between language technologies and multilingual information retrieval.

Research Methodology

This study employs a quantitative research design to examine the impact of AI-driven language technologies on multilingual information retrieval within ASEAN humanities. Established instruments were utilised to assess variable relationships, following empirical analysis

recommendations from prior research. The measurement instruments were adapted from existing studies (An et al., 2023; Safdar et al., 2020). Data collection was conducted using a five-point Likert scale (Jebb et al., 2021). AI was evaluated based on its role in multilingual information retrieval, while language technologies were assessed concerning the effectiveness of machine learning and language translation tools in facilitating multilingual information retrieval. Additionally, multilingual information retrieval was measured to determine AI integration in the multilingual contexts of ASEAN countries.

Data were collected from academic staff at four universities in Malaysia, Indonesia, Thailand, and Brunei. These institutions were selected based on accessibility, with academic staff information retrieved from university websites. A convenience sampling method was employed, and respondents were approached in person to complete printed questionnaires. A total of 410 printed questionnaires were distributed. However, preliminary data analysis was conducted to remove responses with missing values or biases, resulting in a final sample of 357 responses. Data analysis was performed using Smart PLS 4, employing Partial Least Squares–Structural Equation Modelling (PLS-SEM). The analysis followed a two-stage approach: measurement model assessment and structural model assessment.

Data Analysis

Demographic information of the participants was collected and is presented in Table 1. The findings indicate that 53% of the participants were teachers, while 47% were students. Regarding AI usage frequency, 32% of participants reported using AI tools daily, 21% used them weekly, 22% used them monthly, 17% used them twice a week, and 8% used AI every second week. These findings provide insights into the extent of AI adoption among the study participants. Additionally, the geographical distribution of participants shows that 10% were from Malaysia, 14% from Brunei, 36% from Indonesia, and 40% from Thailand.

Table 1: Demographics.

Variable	Level	Counts	Proportion
Occupation	Teachers	188	53%
	Students	169	47%
AI Tools Usage	Daily	114	32%
	Weekly	76	21%
	Monthly	80	22%
	Twice a Week	60	17%
	Every Second Week	27	8%
	Malaysia	35	10%
Country	Brunei	51	14%
	Indonesia	127	36%
	Thailand	144	40%

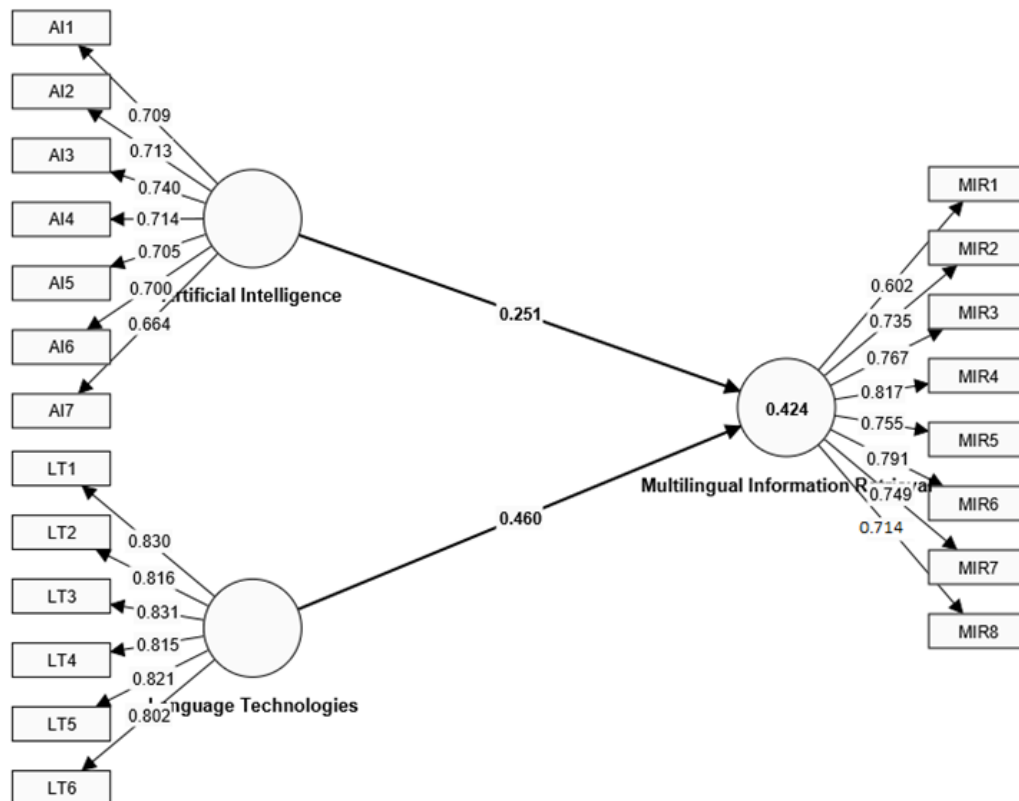


Figure 1: Measurement Model.

The measurement model assessment was conducted to evaluate the reliability of individual items. Factor loadings were examined to determine whether the instruments effectively measured the variables, with all factor loading values exceeding 0.60 and achieving statistical significance. Additionally, composite reliability and Cronbach's alpha were analysed to assess the overall reliability of the instruments. The results indicated that all values for composite reliability and Cronbach's alpha were above 0.70, confirming the reliability of the model (Hair et al., 2011). Furthermore, the study examined the average variance extracted (AVE) to assess the variance explained by the constructs. The AVE values exceeded the threshold

of 0.50, demonstrating satisfactory convergent validity (Hair et al., 2011). These findings indicate that the data is both reliable and valid. The results are presented in Table 2 and Figure 1. In addition to this, the study proceeded with an assessment of discriminant validity to ensure that the constructs measured distinct variables. Discriminant validity was evaluated using the Heterotrait-Monotrait (HTMT) method, as recommended by Henseler et al. (2015). According to this criterion, HTMT values below 0.90 indicate that the constructs are sufficiently distinct. The results, presented in Table 3, confirmed that all HTMT values were below the 0.90 threshold, thereby establishing significant discriminant validity for the research data.

Table 2: Convergent Validity.

Construct	Items	Factor Loadings	Cronbach's Alpha	Composite Reliability	Average Variance Extracted
Artificial Intelligence	AI1	0.709	0.843	0.852	0.699
	AI2	0.713			
	AI3	0.740			
	AI4	0.714			
	AI5	0.705			
	AI6	0.700			
	AI7	0.664			
Language Technologies	LT1	0.830	0.902	0.902	0.671
	LT2	0.816			
	LT3	0.831			
	LT4	0.815			
	LT5	0.821			
	LT6	0.802			
Multilingual Information Retrieval	MIR1	0.602	0.861	0.901	0.511
	MIR2	0.735			
	MIR3	0.767			
	MIR4	0.817			
	MIR5	0.755			
	MIR6	0.791			
	MIR7	0.749			
	MIR8	0.714			

Table 3: Discriminant Validity.

Construct	Artificial Intelligence	Language Technologies	Multilingual Information Retrieval
Artificial Intelligence			
Language Technologies	0.631		
Multilingual Information Retrieval	0.535	0.618	

The study conducted structural model analysis at a significance level of 0.05 to examine the hypothesised relationships between variables. The results confirmed the acceptance of H1, indicating a significant relationship between artificial intelligence and multilingual information retrieval. Similarly, H2 was also supported, demonstrating a significant relationship between language technologies and multilingual information retrieval. The findings of the structural model analysis are presented in Figure 2

and Table 4. Moreover, the study further examined the predictive relevance of the model to assess its overall predictive power. Predictive relevance, represented by Q^2 , indicates the extent to which independent variables contribute to explaining the dependent variable, with values above 0 considered significant (Hair et al., 2011). The findings presented in Table 5 and Figure 3 confirmed that the model demonstrated significant predictive relevance.

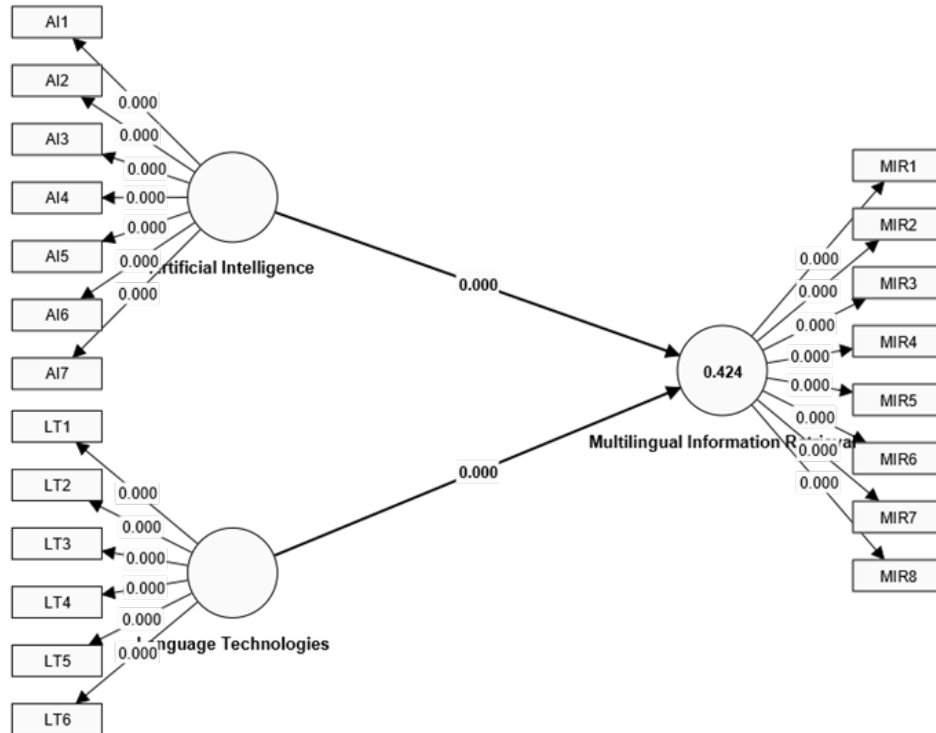


Figure 2: Structural Model.

Table 4: Path Findings.

Paths	Original Sample	Standard Deviation	T Statistics	P Values
Artificial Intelligence -> Multilingual Information Retrieval	0.251	0.048	5.218	0.000
Language Technologies -> Multilingual Information Retrieval	0.460	0.043	10.649	0.000

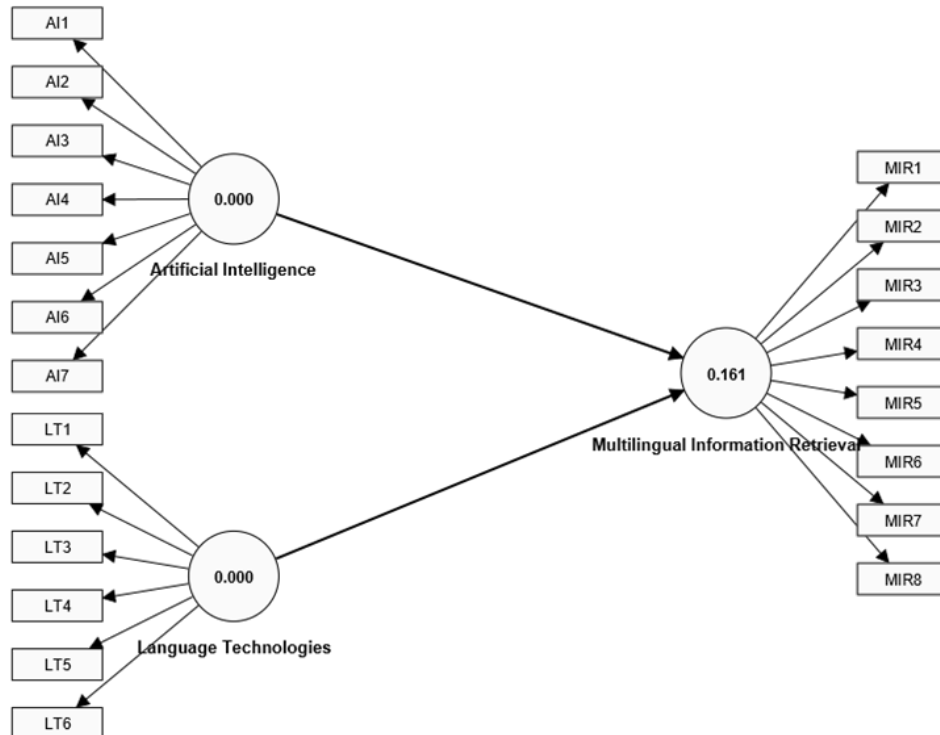


Figure 3: Predictive Relevance.

Table 5: Predictive Relevance.

Construct	SSO	SSE	Q ² (=1-SSE/SSO)
Artificial Intelligence	2499	2499	0
Language Technologies	2142	2142	0
Multilingual Information Retrieval	2856	2395.751	0.161

Discussion

The findings of this research are based on the empirical analysis of the data. The study established a relationship between artificial intelligence (AI) and multilingual information retrieval, which was compared with previous literature. According to Chonka et al. (2023), advancements in AI technology have facilitated information retrieval from multiple sources by employing language modules to comprehend linguistic nuances across various cultures and contexts. Khan et al. (2024) emphasised that a thorough understanding of AI-based systems is essential to enhance information provision and support resources. Similarly, Jantscher et al. (2023) reported that AI significantly benefits information management systems by categorising practical information into distinct sections based on AI-driven insights, thereby ensuring efficient and precise information delivery (Alsalmi, 2021). Furthermore, Ali and Farha (2025) highlighted the necessity for AI developers to understand diverse cultural values to accurately interpret information and queries in multiple languages, thereby improving AI's transparency and information management capabilities.

AI-driven technologies play a crucial role in facilitating information retrieval and document access (Efimov et al., 2023). According to Khurana et al. (2023), these technologies enable academics and students to acquire knowledge by accessing information from diverse sources. AI systems possess the capability to interpret user queries and structure retrieval processes accordingly. Huang and Huang (2024) asserted that the AI era is enhanced through the training of systems that optimise knowledge retrieval. Ahmad et al. (2023) further highlighted AI's essential role in sustaining information management systems by ensuring accurate query interpretation and precise information delivery. Moreover, AI tools must provide exact and relevant information (Litschko et al., 2022). Consequently, extensive training for AI programmes is required to enhance retrieval accuracy and subject-based categorisation, thereby ensuring transparency in disseminated data (Lee and Kim, 2023). Effective information management necessitates AI utilisation, particularly when handling multilingual data. To enhance AI productivity, engineers must refine its efficiency across various languages and cultural contexts (Fan et al., 2022),

enabling AI tools to comprehend queries accurately and deliver precise responses.

Furthermore, the study established a relationship between language technologies and multilingual information retrieval, which was also compared with previous studies. According to Kim et al. (2024), training language models for enhanced multilingual information retrieval improves the productivity of organisations and individuals. Siciliani et al. (2024) noted that in institutions where knowledge is accessed from diverse linguistic sources, language technologies play a crucial role in information sharing and complex data management. Yang et al. (2022) emphasised that AI-driven programming and information retrieval are vital for fostering comprehension and development in multilingual environments. Similarly, Safdar et al. (2020) suggested that AI-powered information retrieval systems significantly enhance efficiency and productivity.

Like AI, language technologies, particularly machine learning mechanisms, are crucial for information retrieval (de Oliveira et al., 2023). These technologies effectively categorise and interpret user queries. Dash (2024) argued that improving linguistic capabilities is fundamental for AI development, enabling machines to generate accurate responses tailored to user needs. As language-based technologies become integral to modern workplaces, they facilitate seamless information access (Ayotunde et al., 2023). However, Alothman and Wahab Sait (2022) stressed that machine learning methodologies processing multilingual data while incorporating cultural values are essential for AI advancement. Additionally, language technologies contribute significantly to AI systems by ensuring the provision of essential information (Lawrie et al., 2023). Although information retrieval is complex, machine learning simplifies the process by structuring information based on specified queries (Sansone and Sperli, 2022). Barsha and Munshi (2024) underscored the role of AI and machine learning in information exchange and retrieval, facilitating enhanced opportunities for accessing data. Effective information retrieval necessitates users' comprehension and reading skills relevant to query formulation (Zhu et al., 2024). AI substantially improves information exchange by integrating machine learning and advanced information management systems.

Conclusion

This study examines the digitisation of the humanities in ASEAN, facilitated by AI and language technologies, as both an opportunity and a challenge. AI has significantly enhanced multilingual information retrieval across the region. However, disparities in implementation persist due to variations in digital infrastructure, linguistic diversity, and the availability of high-quality datasets. Machine translation and natural language processing models, driven by AI, have expanded access to educational and cultural resources, fostering a more inclusive digital landscape. Nevertheless, the effectiveness of these technologies depends on the quality of linguistic datasets and the adaptability of AI models to regional dialects and cultural contexts. Within ASEAN, an analysis of AI adoption reveals a growing divide between technologically advanced nations such as Singapore and Malaysia and those with insufficient digital infrastructure, including Myanmar and Cambodia. These disparities risk exacerbating inequalities in access to information and educational opportunities, particularly where investments in digital development are inadequate or misaligned.

Implications

This study has significant implications for practitioners, stakeholders, and policymakers. It highlights the limitations of AI-driven NLP in text classification and retrieval, particularly due to inconsistencies in transliteration and language structures that hinder AI's ability to process multilingual content effectively. While advancements in deep learning have substantially enhanced AI performance, the availability of large, high-quality datasets remains crucial for further improvement. Additionally, practitioners and stakeholders must recognise that the digital divide in AI infrastructure exacerbates disparities, preventing nations with limited computational resources from fully benefiting from AI technologies. Addressing these challenges requires a coordinated effort to expand datasets, foster regional collaboration, and fine-tune AI models to accommodate ASEAN's linguistic diversity. Moreover, as AI becomes increasingly embedded in multilingual information retrieval, ethical concerns must be considered. Stakeholders and practitioners must ensure that AI models are designed to minimise the risks of cultural homogenisation and the marginalisation of indigenous languages. AI is transforming access to multilingual academic materials and personalised learning tools, offering new opportunities for students and researchers. However, ensuring equitable access remains a challenge,

particularly for remote and underserved communities that lack the necessary digital infrastructure. To prevent AI-driven resources from deepening existing educational inequalities, governments, educational institutions, and technology developers must collaborate to promote inclusive access to AI technologies.

Future efforts to ensure the ethical use of AI should focus on expanding diverse linguistic datasets, improving AI's contextual understanding, and enhancing its ability to process regional dialects. As ASEAN nations continue to benefit from AI advancements, increased investments in AI infrastructure, including cloud-based NLP platforms, will be essential. To standardise AI-driven multilingual information retrieval while preserving linguistic and cultural diversity, regional cooperation and policy frameworks must be established. While AI has the potential to bridge language barriers, enhance educational access, and protect cultural heritage, it must adhere to ethical guidelines and be adapted to the specific linguistic contexts of ASEAN. Governments within ASEAN should integrate universities and linguists into AI development initiatives to enhance language learning mechanisms and information retrieval capabilities. Policy reforms at the governmental level can facilitate the development of AI-driven information retrieval frameworks, ensuring that AI systems are effectively utilised for analysing and disseminating information. Regulatory measures are necessary to optimise AI's role in information retrieval while preserving linguistic and cultural nuances. Additionally, universities must contribute to AI customisation and machine learning advancements that support ASEAN's multilingual landscape, enabling seamless information exchange. By implementing these strategies, policymakers and practitioners can enhance language accessibility and information management across ASEAN, ultimately benefiting the public.

Future Directions

Given that this study collected data from only four universities within ASEAN countries, future research should address this limitation by expanding data collection to the remaining ASEAN nations. Such an approach would enhance the generalisability of the findings across the region. Additionally, this study examines the impact of AI and language technologies on multilingual information retrieval. Future research should explore the role of information retrieval capabilities as a mediating factor between AI, language technologies, and multilingual information retrieval. Investigating this relationship would offer a significant contribution to the existing body

of knowledge. Furthermore, this study employed cross-sectional data collected at a single point in time. While cross-sectional data is valuable in contemporary research due to the ease of accessing respondents, future studies should consider adopting a longitudinal research design. A longitudinal approach would provide deeper insights into the evolving dynamics of AI-driven information retrieval systems over time. Addressing these recommendations in future research will not only strengthen the findings but also contribute to bridging existing gaps in the literature.

Funding Statement

This project, titled “**Constructing ASEAN Sustainable Values Based on Corpus and Critical Discourse Analysis**,” is supported by the 2024 Annual Project of the Belt and Road Country and Regional Research Center of the National Ethnic Affairs Commission (NEAC), Southeast Asia Research Center of Sichuan Normal University (Project No. **2024DNYZC018**). The funding from NEAC’s Belt and Road Research Program supports data collection, analysis, and academic collaboration, contributing to ASEAN-China cooperation and the Belt and Road Initiative.

References

- Ahmad, P. N., Liu, Y., Khan, K., Jiang, T. and Burhan, U. (2023). BIR: Biomedical Information Retrieval System for Cancer Treatment in Electronic Health Record Using Transformers. *Sensors*, 23(23): 9355. <https://doi.org/10.3390/s23239355>
- Ali, L. H. and Farha, N. D. (2025). Bilingualism and Artificial Intelligence: A Sociolinguistic Approach to Arab Heritage. In: A. Abdelgawad, A. Jamil, & A. A. Hameed (Eds.), *Intelligent Systems, Blockchain, and Communication Technologies*. Springer Nature Switzerland, pp. 619-628. https://doi.org/10.1007/978-3-031-82377-0_50
- Alon, L. and Krtalić, M. (2025). “I wish I could use any language as it comes to mind”: User experience in digital platforms in the context of multilingual personal information management. *Journal of the Association for Information Science and Technology*, 76(4): 686-702. <https://doi.org/10.1002/asi.24964>
- Allothman, A. F. and Wahab Sait, A. R. (2022). Managing and Retrieving Bilingual Documents Using Artificial Intelligence-Based Ontological Framework. *Computational Intelligence and Neuroscience*, 2022(1): 4636931. <https://doi.org/10.1155/2022/4636931>
- Alsalmi, H. M. (2021). Information-seeking in multilingual digital libraries. *Library Hi Tech*, 39(1): 80-100. <https://doi.org/10.1108/LHT-06-2019-0119>
- Alsayed, A., Arif, M., Qadah, T. M. and Alotaibi, S. (2023). A Systematic Literature Review on Using the Encoder-Decoder Models for Image Captioning in English and Arabic Languages. *Applied Sciences*, 13(19): 10894. <https://doi.org/10.3390/app131910894>
- An, X., Chai, C. S., Li, Y., Zhou, Y. and Yang, B. (2023). Modeling students’ perceptions of artificial intelligence assisted language learning. *Computer Assisted Language Learning*: 1-22. <https://doi.org/10.1080/09588221.2023.2246519>
- Ayotunde, O. O., Jamil, D. I. and Cavus, N. (2023). The Impact of Artificial Intelligence in Foreign Language Learning Using Learning Management Systems: a Systematic Literature Review. *Information Technologies and Learning Tools*, 95(3): 215-228. <https://doi.org/10.33407/itlt.v95i3.5233>
- Barsha, S. and Munshi, S. A. (2024). Implementing artificial intelligence in library services: a review of current prospects and challenges of developing countries. *Library Hi Tech News*, 41(1): 7-10. <https://doi.org/10.1108/LHTN-07-2023-0126>
- Chonka, P., Diepeveen, S. and Haile, Y. (2023). Algorithmic power and African indigenous languages: search engine autocomplete and the global multilingual Internet. *Media, Culture & Society*, 45(2): 246-265. <https://doi.org/10.1177/01634437221104705>
- Dash, N. S. (2024). Corpus Linguistics and Language Technology. In: C. Sin-Wai, M. Kin-Wah, & L. S. Ming (Eds.), *Routledge Encyclopedia of Technology and the Humanities*. Routledge, pp. 219-246. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781003195269-13>
- de Oliveira, L. L., Vargas, D. S., Alexandre, A. M. A., Cordeiro, F. C., Gomes, D. d. S. M., Rodrigues, M. d. C. *et al.* (2023). Evaluating and mitigating the impact of OCR errors on information retrieval. *International Journal on Digital Libraries*, 24(1): 45-62. <https://doi.org/10.1007/s00799-023-00345-6>
- Efimov, P., Boytsov, L., Arslanova, E. and Braslavski, P. (2023). The Impact of Cross-Lingual Adjustment of Contextual Word Representations on Zero-Shot Transfer. In: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, & A. Caputo (Eds.), *Advances in Information*

- Retrieval*. Springer Nature Switzerland, pp. 51-67. https://doi.org/10.1007/978-3-031-28241-6_4
- Fan, Y., Xie, X., Cai, Y., Chen, J., Ma, X., Li, X. *et al.* (2022). Pre-training methods in information retrieval. *Foundations and Trends® in Information Retrieval*, 16(3): 178-317. <https://doi.org/10.1561/15000000100>
- Hair, J. F., Ringle, C. M. and Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, 19(2): 139-152. <https://doi.org/10.2753/MTP1069-6679190202>
- Han, L., Gladkoff, S., Erofeev, G., Sorokina, I., Galiano, B. and Nenadic, G. (2024). Neural machine translation of clinical text: an empirical investigation into multilingual pre-trained language models and transfer-learning. *Frontiers in Digital Health*, 6: 1211564. <https://doi.org/10.3389/fdgth.2024.1211564>
- Heng, S., Tsilonis, K., Scharff, C. and Wautelet, Y. (2022). Understanding AI ecosystems in the Global South: The cases of Senegal and Cambodia. *International Journal of Information Management*, 64: 102454. <https://doi.org/10.1016/j.ijinfomgt.2021.102454>
- Henseler, J., Ringle, C. M. and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1): 115-135. <https://doi.org/10.1007/s11747-014-0403-8>
- Horna-Saldaña, C. J., Perez Perez, J. E. and Toro Galeano, M. L. (2025). Artificial intelligence in the preservation of native languages and bridging the information access gap for indigenous peoples. *Journal of Enabling Technologies*, 19(1): 63-75. <https://doi.org/10.1108/JET-09-2024-0063>
- Huang, Y. and Huang, J. X. (2024). Exploring ChatGPT for next-generation information retrieval: Opportunities and challenges. *Web Intelligence*, 22(1): 31-44. <https://doi.org/10.3233/WEB-230363>
- Jantscher, M., Gunzer, F., Kern, R., Hassler, E., Tschauner, S. and Reishofer, G. (2023). Information extraction from German radiological reports for general clinical text and language understanding. *Scientific Reports*, 13(1): 2353. <https://doi.org/10.1038/s41598-023-29323-3>
- Jebb, A. T., Ng, V. and Tay, L. (2021). A Review of Key Likert Scale Development Advances: 1995–2019. *Frontiers in psychology*, 12: 637547. <https://doi.org/10.3389/fpsyg.2021.637547>
- Khan, R., Gupta, N., Sinhababu, A. and Chakravarty, R. (2024). Impact of Conversational and Generative AI Systems on Libraries: A Use Case Large Language Model (LLM). *Science & Technology Libraries*, 43(4): 319-333. <https://doi.org/10.1080/0194262X.2023.2254814>
- Khan, W., Daud, A., Khan, K., Muhammad, S. and Haq, R. (2023). Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends. *Natural Language Processing Journal*, 4: 100026. <https://doi.org/10.1016/j.nlp.2023.100026>
- Khurana, D., Koli, A., Khatter, K. and Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*, 82(3): 3713-3744. <https://doi.org/10.1007/s11042-022-13428-4>
- Kim, H.-W., Shin, D.-H., Kim, J., Lee, G.-H. and Cho, J. W. (2024). Assessing the performance of ChatGPT's responses to questions related to epilepsy: A cross-sectional study on natural language processing and medical information retrieval. *Seizure: European Journal of Epilepsy*, 114: 1-8. <https://doi.org/10.1016/j.seizure.2023.11.013>
- Lawrie, D., Yang, E., Oard, D. W. and Mayfield, J. (2023). Neural Approaches to Multilingual Information Retrieval. In: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, & A. Caputo (Eds.), *Advances in Information Retrieval*. Springer Nature Switzerland, pp. 521-536. https://doi.org/10.1007/978-3-031-28244-7_33
- Lee, C.-H., Wang, C., Fan, X., Li, F. and Chen, C.-H. (2023). Artificial intelligence-enabled digital transformation in elderly healthcare field: Scoping review. *Advanced Engineering Informatics*, 55: 101874. <https://doi.org/10.1016/j.aei.2023.101874>
- Lee, J. and Kim, M. (2023). ESG information extraction with cross-sectoral and multi-source adaptation based on domain-tuned language models. *Expert Systems with Applications*, 221: 119726. <https://doi.org/10.1016/j.eswa.2023.119726>
- Lin, Y. and Yu, Z. (2025). Learner Perceptions of Artificial Intelligence-Generated Pedagogical Agents in Language Learning Videos: Embodiment Effects on Technology Acceptance. *International Journal of Human-Computer Interaction*, 41(2): 1606-1627. <https://doi.org/10.1080/10447318.2024.2359222>

- Litschko, R., Vulić, I., Ponzetto, S. P. and Glavaš, G. (2022). On cross-lingual retrieval with multilingual text encoders. *Information Retrieval Journal*, 25(2): 149-183. <https://doi.org/10.1007/s10791-022-09406-x>
- Liu, G. and Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 18(2): 125-138. <https://doi.org/10.1080/17501229.2023.2240316>
- Mustapa, K., Gagaramusu, Y. B. M., Palandi, E. H., Syakhrani, A. W. and Towpek, H. (2023). Technology-enhanced education: Nurturing the digital generation-experiences in Islamic schools in Indonesia. *International Journal of Teaching and Learning*, 1(1): 16-40. <https://injournal.org/index.php/12/article/view/2>
- Nzomo, P., McKenzie, P., Ajiferuke, I. and Vaughan, L. (2021). Towards a Definition of Multilingual Information Literacy (MLIL): An Essential Skill for the 21st Century. *Journal of Library Administration*, 61(7): 897-920. <https://doi.org/10.1080/01930826.2021.1972737>
- Safdar, Z., Bajwa, R. S., Hussain, S., Abdullah, H. B., Safdar, K. and Draz, U. (2020). The role of Roman Urdu in multilingual information retrieval: A regional study. *The Journal of Academic Librarianship*, 46(6): 102258. <https://doi.org/10.1016/j.acalib.2020.102258>
- Sansone, C. and Sperli, G. (2022). Legal Information Retrieval systems: State-of-the-art and open issues. *Information Systems*, 106: 101967. <https://doi.org/10.1016/j.is.2021.101967>
- Siciliani, L., Ghizzota, E., Basile, P. and Lops, P. (2024). OIE4PA: open information extraction for the public administration. *Journal of Intelligent Information Systems*, 62(1): 273-294. <https://doi.org/10.1007/s10844-023-00814-z>
- Yang, Y., Wu, Z., Yang, Y., Lian, S., Guo, F. and Wang, Z. (2022). A Survey of Information Extraction Based on Deep Learning. *Applied Sciences*, 12(19): 9691. <https://doi.org/10.3390/app12199691>
- Zhu, J., Wu, J., Luo, X. and Liu, J. (2024). Semantic matching based legal information retrieval system for COVID-19 pandemic. *Artificial Intelligence and Law*, 32(2): 397-426. <https://doi.org/10.1007/s10506-023-09354-x>



Qi Zhu holds a Ph.D. from Nanjing Normal University, specializing in Teaching Chinese as a Foreign Language. Her research focuses on classroom instruction in teaching Chinese as a foreign language, learning strategies, and region-specific approaches to language education.



Zijun Shen, PhD in multicultural studies and sociolinguistics. Zijun Shen's research interests include in intercultural communication and cultural studies between Japan and China, and sociolinguistics in ASEAN studies. Zijun Shen also serves as the editorial member and reviewer of the World Journal of English Language.