Mahrajan Arabi
Proceeding International Conference in Arabic Festival (INCAFA) 2026

ARABIC AS A GLOBAL LANGUAGE IN THE DIGITAL ERA: A CRITICAL ANALYSIS OF ARTIFICIAL INTELLIGENCE-BASED TRANSLATION AND LANGUAGE AUTHORITY

Aufia Aisa*[1], Imam Asrori*[2], Siswi Yohana*[3]
Universitas Negeri Malang[123]
aufia.aisa.2502319@students.um.ac.id, imam.asrori.fs@um.ac.id,
siswi.yohana.2502319@students.um.ac.id

Abstract

This study examines Arabic as a global language in the digital era, focusing on AI-based translation and linguistic authority. Using a qualitative conceptual approach, the study integrates perspectives from sociolinguistics and computer-mediated communication. The findings indicate that AI translation expands the global reach of Arabic but also creates challenges related to contextual meaning, linguistic accuracy, and language authority. The study emphasizes the importance of balancing technological innovation with established linguistic scholarship.

Keyword: Arabic language; digital era; artificial intelligence; translation; language authority

مستخلص البحث

تهدف هذه الدراسة إلى تحليل اللغة العربية بوصفها لغة عالمية في العصر الرقمي مع التركيز على الترجمة القائمة على الذكاء الاصطناعي والسلطة اللغوية. اعتمدت الدراسة على المنهج الكيفي المفاهيمي من خلال توظيف علم اللغة الاجتماعي والاتصال الحاسوبي. وأظهرت النتائج أن الترجمة بالذكاء الاصطناعي توسّع انتشار اللغة العربية عالميًا، لكنها تثير تحديات تتعلق بالدقة اللغوية والسياق والسلطة اللغوية. وتؤكد الدراسة أهمية التوازن بين التطور التكنولوجي والمرجعية اللغوية الراسخة.

كلمات أساسية: اللغة العربية؛ العصر الرقمي؛ الذكاء الاصطناعي؛ الترجمة؛ السلطة اللغوية

The development of digital technology has brought significant changes to global communication patterns, including the use and dissemination of language. Arabic, as one of the international languages with strong religious, historical, and cultural dimensions, is

currently facing new dynamics within the digital sphere. This transformation is marked by the increasing use of social media, online communication platforms, and artificial intelligence (AI) technologies that enable Arabic to be accessed and used more broadly by global communities (Warschauer, 2011). In this context, Arabic is no longer positioned solely as a religious and cultural language, but also as part of technology-based global communication.

The advancement of AI-based technologies, particularly in machine translation, has accelerated the spread of Arabic across linguistic and geographical boundaries. Neural machine translation systems are capable of producing rapid and real-time translations, thereby facilitating cross-language communication. The presence of this technology provides significant opportunities to improve the accessibility of Arabic in education, academia, and international communication (Chiu, 2024). Furthermore, the emergence of generative AI technologies such as ChatGPT and Midjourney has transformed educational practices, policies, and research directions in the digital era (Chiu, 2025). Nevertheless, the use of AI in Arabic translation also raises various issues related to linguistic accuracy, contextual meaning, and the validity of language interpretation.

From a sociolinguistic perspective, these changes affect the dynamics of Arabic language use, particularly concerning the phenomenon of diglossia between *fusha* and *amiyah*. Digital environments encourage more flexible, hybrid, and informal forms of communication, causing the boundaries between formal and informal varieties to become increasingly blurred (Bassiouney, 2020). This phenomenon is further supported by the findings of Alfaisal and Aljanada (2019), which indicate that diglossia in Arabic continues to evolve alongside the growing intensity of social interaction in digital spaces. This condition demonstrates that Arabic is undergoing transformation not only in technological aspects but also in social and cultural dimensions.

In addition, the development of AI as a linguistic mediator has introduced new concerns regarding language authority. Authority that was previously dominated by scholars, linguists, and native speakers is now gradually shifting toward systems based on algorithms and digital data (Fairclough, 2010; Van Dijk, 2015). In this context, technology functions not merely as a communication tool but also as an instrument that shapes power relations and the production of linguistic knowledge. Chomsky, Roberts, and Watumull (2023) argue that AI systems do not genuinely understand language but merely predict patterns based on available data. This criticism highlights the tension between human linguistic authority and algorithmic authority.

On the other hand, the development of digital literacy and multimodal communication also influences the way Arabic is used and understood. Language in digital environments is no longer presented only in textual forms, but also through images, videos, emojis, and other visual symbols (Kress, 2010). This transformation indicates that modern communication is inherently multimodal and requires more complex digital literacy skills. Li

(2024) found that digital reading media significantly influence users' comprehension patterns and interactions with texts. In addition, Sinaga et al. (2024) emphasize that digital literacy can foster creativity and communication skills in modern educational environments.

Studies on Arabic in the digital era have been widely conducted; however, most of them focus primarily on technological effectiveness and the development of translation systems. Research critically examining the relationship between AI, linguistic authority, and the position of Arabic as a global language remains relatively limited. Previous studies have mainly emphasized technical aspects of digital learning and educational technology implementation (Huang et al., 2019; Tlili et al., 2023), while the epistemological and ideological dimensions of Arabic within the context of AI have not been extensively explored. In fact, the rise of generative AI has sparked serious debates regarding the validity of linguistic knowledge and language representation in digital spaces.

Moreover, the development of digital corpora and computational linguistics has also influenced modern Arabic language studies. Alfaifi and Atwell (2016) explain that Arabic learner corpora have become an essential component in the development of AI and machine learning systems for Arabic. This aligns with McEnery and Hardie's (2012) argument that corpus linguistics plays a crucial role in understanding language patterns based on digital data. Therefore, the advancement of AI in Arabic cannot be separated from the processes of collecting, managing, and interpreting large-scale linguistic data.

Based on this background, this study aims to critically analyze the position of Arabic as a global language in the digital era, focusing on AI-based translation and the shifting dynamics of language authority. This study employs a qualitative conceptual approach by integrating perspectives from computer-mediated communication, sociolinguistics, language ideology, and digital epistemology. Through this approach, the study is expected to provide a more comprehensive understanding of the transformation of Arabic within the global digital ecosystem and its implications for contemporary linguistic development.

Methods

This study employed a qualitative conceptual approach with a systematic literature review and critical analysis design. The study aimed to critically examine the position of Arabic as a global language in the digital era, particularly in relation to AI-based translation and the transformation of linguistic authority. The qualitative conceptual method was selected because it allows the researcher to explore theoretical perspectives, synthesize previous studies, and critically interpret contemporary linguistic phenomena in digital contexts.

The primary data sources consisted of international peer-reviewed journal articles, academic books, conference proceedings, and official reports related to artificial intelligence, Arabic sociolinguistics, digital communication, and language ideology. The literature was obtained from reputable academic databases such as Scopus, Google Scholar, Springer, Routledge, and UNESCO publications. The selection of references followed several inclusion criteria: (1) publications relevant to Arabic language studies, AI-based translation, digital

communication, and sociolinguistics; (2) publications indexed in reputable databases; and (3) sources published mainly within the last ten years to ensure the relevance and currency of the discussion.

Data collection was conducted through systematic documentation and literature tracing using keywords such as "Arabic language in the digital era," "AI-based translation," "language authority," "digital literacy," "computer-mediated communication," and "Arabic sociolinguistics." The collected literature was then classified based on thematic relevance and theoretical contribution. In addition, supporting materials such as digital communication studies, corpus linguistics research, and AI-related educational studies were included to strengthen the interdisciplinary framework of the research.

The data analysis process employed content analysis and critical synthesis techniques (Krippendorff, 2019). The researcher identified major concepts, categorized themes, and interpreted findings using theoretical frameworks from computer-mediated communication (Herring, 1996), sociolinguistics (Bassiouney, 2020), critical discourse analysis (Fairclough, 2010; Van Dijk, 2015), multimodality (Kress, 2010), and connectivism (Siemens, 2005). This analytical approach enabled a comprehensive understanding of how AI-based translation technologies reshape Arabic language practices, linguistic authority, and epistemological perspectives in the digital era.

The research instruments used in this study consisted of literature review matrices, thematic coding tables, and conceptual mapping frameworks designed to organize and synthesize the collected data systematically. These instruments facilitated the identification of relationships between technological development, language transformation, and digital communication practices. Through this methodological framework, the study provides a critical and interdisciplinary analysis of Arabic language transformation within the contemporary global digital ecosystem.

Results

The findings of this study indicate that the development of digital technology and artificial intelligence (AI) has significantly transformed the position and use of Arabic in global communication. Based on the literature analysis, this transformation can be observed through the increasing use of digital platforms, social media, and AI-based translation systems that enable Arabic to be used more widely in cross-cultural and cross-linguistic interactions.

The findings reveal that AI-based translation technology contributes substantially to improving the accessibility of the Arabic language. Neural machine translation systems enable rapid and real-time translation processes, making it easier for users to understand Arabic texts. In addition, this technology supports the global dissemination of Arabic academic, religious, and cultural content. Nevertheless, the study also found that AI systems still face limitations in understanding cultural contexts, pragmatic nuances, and contextual meanings in Arabic.

This study also found changes in Arabic communication patterns within digital environments. The use of social media and online communication platforms encourages the emergence of more flexible and hybrid language forms. The boundaries between the formal

variety (*fusha*) and the informal variety (*amiyah*) are becoming increasingly blurred due to the intensity of digital interaction. This phenomenon demonstrates that Arabic is undergoing socio-linguistic transformation influenced by the development of modern communication technologies.

Furthermore, the findings indicate a shift in language authority in the digital era. Linguistic authority, which was previously dominated by scholars, linguists, and academic institutions, is gradually shifting toward algorithm-based systems and digital communities. In this context, AI functions not only as a translation tool but also as a linguistic mediator that influences how users understand and produce Arabic. This shift affects the process of validating meaning and interpreting language in digital spaces.

Based on the synthesis of the literature, this study also found that the development of AI in Arabic is closely related to the use of digital corpora and large-scale linguistic data. AI systems operate based on data patterns available in language corpora; therefore, the quality of translation outputs is highly influenced by the linguistic representation contained in the data. This condition may lead to linguistic bias, particularly because most digital corpora predominantly represent *fusha* rather than local variations or *amiyah*.

To clarify the analysis results, the following table presents the transformation of Arabic in the digital era.

Table 1. Transformation of Arabic Language in the Digital Era

Aspect	Research Findings	Impact
Digital Technology	Increased use of digital media and online platforms	Global dissemination of Arabic
AI Translation	Faster and more accessible translation	Improved communication accessibility
Digital Diglossia	Blurred boundaries between *fusha* and *amiyah*	Emergence of hybrid language forms
Language Authority	Shift from linguists to algorithms	Changes in linguistic legitimacy
Digital Corpora	Dominance of *fusha* data in AI systems	Potential linguistic bias

Source: Synthesized by the author from various literature sources.

The findings demonstrate that digital technology and AI play a dual role in the development of Arabic. On the one hand, technology expands the global reach and use of Arabic. On the other hand, these developments also create new challenges related to linguistic accuracy, language standardization, and epistemological validity in the use of Arabic in the digital era.

Discussions

The findings of this study indicate that the development of digital technology and artificial intelligence has significantly transformed the position of Arabic in global communication. This transformation demonstrates that Arabic no longer develops solely within traditional spaces but also within a digital ecosystem influenced by social media, online platforms, and algorithm-based systems. These findings are consistent with the theory of computer-mediated communication, which emphasizes that digital technology shapes communication patterns, social interaction, and linguistic practices in new ways (Herring, 1996; Warschauer, 2011). In this context, Arabic has expanded its function as a global language integrated with digital technology.

This study found that AI-based translation contributes significantly to the accessibility of Arabic at the international level. These findings support Chiu's (2024) study, which states that generative AI has accelerated the transformation of educational practices and digital communication. Technologies such as ChatGPT enable non-native speakers to access Arabic texts more easily and quickly. In addition, UNESCO (2023) emphasizes that AI has substantial potential to expand access to education and cross-linguistic knowledge. Nevertheless, the findings of this study indicate that increased accessibility is not always accompanied by a deeper quality of language understanding.

From a linguistic perspective, the limitations of AI in understanding cultural contexts and pragmatic meanings indicate that technology has not been able to fully replace human linguistic competence. These findings reinforce the epistemological criticism proposed by Chomsky, Roberts, and Watumull (2023), who argue that AI does not genuinely understand language but merely predicts patterns based on statistical data. Therefore, although AI excels in speed and efficiency, the validity of language interpretation still depends on human ability to understand the social, cultural, and ideological contexts of Arabic.

The phenomenon of digital diglossia identified in this study also demonstrates changes in the communication structure of Arabic in the digital era. The boundaries between *fusha* and *amiyah* have become increasingly flexible due to the use of social media and informal online communication. These findings are in line with the studies of Bassiouney (2020) and Alfaisal and Aljanada (2019), which reveal that digital spaces encourage the emergence of more hybrid and fluid forms of Arabic. From a sociolinguistic perspective, this phenomenon indicates that language is not static but continuously evolves following social and technological changes.

Furthermore, the development of multimodal communication has strengthened the transformation of Arabic within digital environments. Language is no longer presented solely in textual form but is also combined with visual symbols, images, videos, and emojis (Kress, 2010). This condition demonstrates that modern communication practices require more complex digital literacy skills. These findings are relevant to Li's (2024) study, which emphasizes that digital media influence users' comprehension patterns and interactions with texts. In the context of Arabic language learning, this transformation highlights the importance of developing multimodal literacy so that users can understand language meanings more critically and contextually.

This study also found a shift in language authority from linguists and academic institutions toward algorithm-based systems and digital communities. From the perspective of Critical Discourse Analysis, this condition indicates that technology is not a neutral tool but is closely related to power and ideology (Fairclough, 2010; Van Dijk, 2015). AI functions as a linguistic mediator that can influence language legitimacy and the way users interpret meaning. This shift reflects a transition from *knowledge authority* toward *algorithmic authority*, in which language validity is increasingly influenced by data-driven digital systems.

From the perspective of digital epistemology, the findings of this study contribute significantly to understanding the relationship between language, technology, and knowledge production. AI systems operate based on digital corpora and large-scale linguistic data, meaning that language representation is highly influenced by the quality and distribution of the data used (McEnery & Hardie, 2012). In the context of Arabic, the dominance of *fusha* corpora in AI systems may lead to language homogenization and reduce the representation of local variations or *amiyah*. This condition demonstrates that digital technology can influence linguistic knowledge structures and cultural representation within Arabic.

In the field of education, the findings of this study indicate that the integration of AI into Arabic language learning has two interconnected dimensions. On the one hand, technology can improve learning accessibility, flexibility, and communication efficiency. This finding aligns with the theory of connectivism, which emphasizes that knowledge develops through digital networks and information connectivity (Siemens, 2005). On the other hand, excessive dependence on AI may reduce users' critical thinking abilities and the depth of linguistic understanding. Therefore, the use of AI in Arabic language learning should be balanced with the strengthening of critical literacy and contextual linguistic competence.

Overall, this study demonstrates that the transformation of Arabic in the digital era is a multidimensional phenomenon involving linguistic, technological, social, and epistemological aspects simultaneously. Arabic as a global language faces not only opportunities to expand international communication but also challenges in maintaining linguistic authority, meaning validity, and language diversity. Therefore, an integrative and critical approach is necessary to ensure that technological development remains balanced with the preservation of the linguistic and cultural values of Arabic.

Conclusions

This study demonstrates that the development of digital technology and artificial intelligence has brought significant transformation to the position of Arabic as a global language in the digital era. The presence of AI-based translation has expanded the accessibility and dissemination of Arabic in international communication, education, and digital knowledge production. Nevertheless, these developments also create various challenges related to linguistic accuracy, contextual understanding, epistemological validity, and the shift of language authority from linguists toward algorithm-based systems.

The findings reveal that Arabic is undergoing socio-linguistic transformation marked by increasingly flexible boundaries between *fusha* and *amiyah* in digital communication. In addition, the dominance of digital corpora in AI development may create linguistic bias and language homogenization. In this context, AI functions not only as a technological tool but also as a linguistic and epistemological mediator that influences the production, distribution, and legitimacy of Arabic linguistic knowledge.

This study also emphasizes that the use of AI in Arabic should be balanced with the strengthening of digital literacy, critical understanding, and human linguistic competence. Technology should be positioned as a supporting tool that enhances communication and learning processes rather than fully replacing human linguistic authority. Therefore, an integrative and multidisciplinary approach is necessary to ensure that digital technological development remains aligned with the preservation of the linguistic, cultural, and epistemological values of Arabic amid the rapid flow of digital globalization.

Acknowledgement

The author would like to express sincere gratitude to Universitas Negeri Malang for providing academic support throughout the research and writing process of this article. The author also extends appreciation to all lecturers, academic colleagues, and all parties who contributed valuable insights, support, and intellectual contributions that enabled this study to be completed successfully.

References

Abu ʻAyādah, H. T. (n.d.). *Taʻallum al-lughah al-ʻArabiyyah fī ẓill al-taḥawwul al-raqmī: Al-wāqiʻ wa al-taḥaddiyāt* [Learning Arabic considering digital transformation: Reality and challenges]. Proceedings of Third ICON POSTALL Pascasarjana PBA Fakultas Sastra Universitas Negeri Malang.

Alfaifi, A., & Atwell, E. (2016). Arabic learner corpora: A survey. *International Journal of Speech Technology*, *19*(1), 1–16.

Alfaisal, A., & Aljanada, R. (2019). *Diglossia in Arabic: Views and opinions*.

Bassiouney, R. (2020). *Arabic sociolinguistics: Topics in diglossia, gender, identity, and politics* (2nd ed.). Georgetown University Press.

Chiu, T. K. F. (2024). The impact of generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments, 32*(10), 6187–6203. https://doi.org/10.1080/10494820.2023.2253861

Chiu, T. K. F. (2025). Introduction to transforming K–12 education with artificial intelligence. In *Empowering K–12 education with AI: Preparing for the future of education and work* (Chap. 1). Routledge. https://doi.org/10.4324/9781003498377-1

Chomsky, N., Roberts, I., & Watumull, J. (2023, March 8). The false promise of ChatGPT. *The New York Times*. https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html

Fairclough, N. (2010). *Critical discourse analysis: The critical study of language* (2nd ed.). Routledge.

Herring, S. C. (1996). *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. John Benjamins Publishing.

Huang, R., Spector, J. M., & Yang, J. (2019). *Educational technology: A primer for the 21st century*. Springer.

Khalil, M., Prinsloo, P., & Slade, S. (2022). The use and application of learning theory in learning analytics: A scoping review. *Journal of Learning Analytics*, Advance online publication.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Krashen, S. (2013). *Second language acquisition: Theory, applications, and some conjectures*. Cambridge University Press.

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications. https://doi.org/10.4135/9781071878781

Li, C. (2024). A study on the impact of digital picture book reading on children's reading literacy. *Communications in Humanities Research, 34*(1), 5–9. https://doi.org/10.54254/2753-7064/34/20240078

Mahmoud Al-Batal. (n.d.). *Arabic as one language: The integration of dialects in the Arabic language curriculum*.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

Nzenwata, U. J., Barn-Nzekwe, C. L., Ojelabi, E. O., Oduware, O., Atalor, P. E., Yisau, Y., Adeyela, A. T., Nwanguma, E. C., Emokiniovo, E., & Osisanya, O. A. (n.d.). *A systematic review of generative AI in education*.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning, 2*(1), 3–10.

Sinaga, R., Simanjuntak, S., Ginting, M., Tarigan, F., & Pinem, A. (2024). Empowering digital literacy through Wattpad: Bridging creativity, reading, and writing in education. *Jurnal Pengabdian Masyarakat dan Riset Pendidikan, 2*(3), 157–161. https://doi.org/10.31004/jerkin.v2i3.237

Tlili, A., et al. (2023). Is metaverse in education a blessing or a curse: A systematic review. *Smart Learning Environments, 10*(1). https://doi.org/10.1186/s40561-022-00205-x

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Van Dijk, T. A. (2015). *Discourse and power*. Palgrave Macmillan.

Warschauer, M. (2011). *Learning in the cloud: How (and why) to transform schools with digital media*. Teachers College Press.