

# Integrating Artificial Intelligence into Arabic Language Education: Socio-Technical Perspectives from Islamic Institutions in Southeast Asia

<sup>1</sup>Ridha Darmawaty\*, <sup>2</sup>Nur Aina Faqihah, <sup>3</sup>Asep Sunarko, <sup>4</sup>Yusuf Arisandi

Email : <sup>1</sup>[ridhadarmawaty@uin-antasari.ac.id](mailto:ridhadarmawaty@uin-antasari.ac.id)\*, <sup>2</sup>[aufa.hanin@live.iium.edu.my](mailto:aufa.hanin@live.iium.edu.my) ,  
<sup>3</sup>[asepsunarko@unsiq.ac.id](mailto:asepsunarko@unsiq.ac.id) <sup>4</sup>[yusufqarisandi@uiidilwa.ac.id](mailto:yusufqarisandi@uiidilwa.ac.id)

<sup>1</sup>Universitas Islam Negeri Antasari Banjarmasin, Indonesia;

<sup>2</sup>International Islamic University Malaysia, Malaysia; <sup>3</sup>Universitas Sains Al-Quran, Indonesia;

<sup>4</sup>Universitas Islam Internasional Darullughah Wadda'wah, Indonesia

## ABSTRACT

This study examines the integration of artificial intelligence (AI) into Arabic language education in Islamic institutions through a multi-site mixed-methods design across Indonesia and Malaysia. Addressing the limited empirical attention to socio-technical, cultural, and religious dimensions of AI-mediated Arabic learning, the study investigates how AI shapes linguistic development, pedagogical practice, learner autonomy, and institutional decision-making. Data were collected from five educational settings: pesantren in Wonosobo and Bangil, UIN Antasari Banjarmasin, UII Darullughah Wadda'wah Pasuruan, and an Islamic institution in Kedah, Malaysia. Classroom observations, interviews, focus group discussions, surveys, digital artifacts, and learning documents were analyzed through triangulation. The findings show that AI supports Arabic learning most effectively as a micro-scaffolding tool for grammar feedback, vocabulary practice, pronunciation support, error detection, and guided interaction. However, its effectiveness depends on teacher mediation, clear learning objectives, and ethical-pedagogical boundaries. Three main constraints emerged: goal diffusion and over-reliance when AI tasks lack guardrails; uneven digital readiness and infrastructure; and persistent inaccuracies in literary, poetic, and classical Arabic texts requiring *dzawq lughawī*, rhetorical nuance, and religious interpretation. Comparative findings indicate that Indonesian sites adopted AI more progressively through pesantren-based filters, whereas Kedah applied a selective, needs-based approach grounded in traditional learning preferences. Theoretically, this study proposes a socio-technical model that positions AI as a cognitive accelerator within a teacher-led, value-anchored Arabic learning ecosystem. Practically, it offers principles for guardrailed prompting, ethical AI use, teacher capacity-building, and institutionally grounded AI integration.

## Article history

Received 2026-04-22

Revised 2026-05-13

Accepted 2026-06-09

## Keywords

Arabic education;  
artificial  
intelligence; Islamic  
education; mixed  
methods; natural  
language processing

\*Correspondence Author

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



## Introduction

In recent years, Arabic language education has entered a new phase in which artificial intelligence (AI) is no longer viewed merely as a supplementary digital tool, but increasingly as a pedagogical force that shapes how learners access texts, receive feedback, practice communication, and regulate their own learning. This transformation is particularly important in Arabic education because Arabic is not only a linguistic system with complex morphology, syntax, diglossia, and rhetorical traditions, but also a language of religious knowledge, scholarly transmission, and Islamic identity. In pesantren, Islamic universities, and madrasah-based learning communities, Arabic learning therefore involves more than technical language mastery; it also requires adab, teacher guidance, epistemic responsibility, and sensitivity toward classical and religious texts.

Recent studies show that AI and digital learning tools have considerable potential to support Arabic language learning. Generative AI can assist learners in writing, grammar correction, vocabulary development, pronunciation practice, and interactive communication (Alkaabi & Almaamari, 2025; Sahrir et al., 2025; Zubaidi et al., 2025). Studies in Islamic educational contexts have also emphasized the need for pedagogical innovation in Arabic learning, especially in vocabulary acquisition, speaking proficiency, writing development, and pesantren-based curriculum models. For example, recent works in AI-Muhawaroh discuss contemporary Arabic vocabulary acquisition in pesantren, vocabulary learning through Tafsir al-Shawi, Duolingo-based vocabulary acquisition, and typologies of Arabic learning in pesantren contexts. Similarly, studies in Farasyah have highlighted immersive Arabic-Islamic environments, sociolinguistic interaction in Malaysian contexts, and gamified instruction in nahwu–sharaf, suggesting that Arabic learning in Southeast Asia is shaped by linguistic, social, affective, and institutional factors (Kasmaluddin, 2026; Muhith & Batrisya, 2026; Mukhasibi & Wahyudi, 2026).

However, the existing literature still leaves two important gaps. First, studies on AI in Arabic language education tend to emphasize technical potential—such as automated feedback, vocabulary practice, writing assistance, or pronunciation support—without sufficiently explaining how AI should be conceptually positioned within Islamic educational traditions. As a result, the use of theories such as sociolinguistics, psycholinguistics, TPACK, and socio-technical systems often remains declarative rather than integrated into a clear analytical framework. Second, previous research has not sufficiently translated AI-related findings into practical institutional guidance for Islamic Arabic education. There remains limited explanation of how teachers, pesantren, Islamic universities, and madrasah-based institutions

should regulate AI use, prevent over-reliance, maintain teacher authority, and protect the values of *adab*, *sanad*, *niyyah*, and textual authenticity.

This study addresses these gaps by focusing on two aspects: theoretical contribution and practical implications. Theoretically, the study argues that AI integration in Arabic language education should be understood through an integrative socio-technical framework. Sociolinguistics explains how Arabic learning is shaped by institutional norms, language ideology, teacher authority, and community practices (Wardhaugh & Fuller, 2014). Psycholinguistics helps interpret how learners process feedback, internalize grammar, develop confidence, and experience possible cognitive dependency when interacting with AI (Levelt, 2012; X. Wang et al., 2024). TPACK clarifies the relationship between technological tools, pedagogical decisions, and Arabic content knowledge, particularly in grammar, vocabulary, pronunciation, and reading instruction (Mishra & Koehler, 2006; Priyanda et al., 2025). Socio-technical systems theory connects these classroom-level processes with infrastructure, institutional readiness, ethics, and human–technology role division (Baxter & Sommerville, 2011; Chen et al., 2020). Through this integration, AI is conceptualized not as an autonomous teacher, but as a micro-scaffolding tool operating within a teacher-led and value-anchored Arabic learning ecosystem.

Practically, this study responds to the need for institutionally grounded guidance on AI use in Islamic Arabic education. In Indonesian *pesantren* and Islamic universities, AI adoption is increasingly visible, yet it must be filtered through teacher supervision, *pesantren* values, and sensitivity toward *turāth* texts. In Malaysian Islamic educational contexts, particularly in traditional learning environments, AI adoption tends to be more selective because *talaqqī*, face-to-face instruction, and teacher authority remain central to the learning process. These differences make Indonesia and Malaysia a relevant comparative context for examining how AI use is shaped not only by technological access, but also by institutional culture, religious values, and pedagogical judgment. The manuscript itself shows that AI was studied across *pesantren* in Wonosobo and Bangil, UIN Antasari Banjarmasin, UII Darullughah Wadda'wah Pasuruan, and an Islamic institution in Kedah, Malaysia, making the multi-site Indonesia–Malaysia context central to the study.

Based on this background, the study is guided by the following research questions: How can AI integration in Islamic Arabic language education be theoretically explained through an integrative socio-technical framework that connects TPACK, sociolinguistic, psycholinguistic, and Islamic educational values? What practical implications can be drawn for teachers and Islamic institutions in using AI while preserving *adab*, *sanad*, teacher authority, and epistemic integrity in Arabic learning?

The central argument of this article is that AI can support Arabic language learning when it functions as a micro-scaffolding tool for error detection, grammar feedback, vocabulary practice, pronunciation support, and guided interaction. However, AI becomes pedagogically meaningful only when it is embedded within clear learning objectives, teacher mediation, ethical boundaries, and institutional values. In literary, poetic, and classical Arabic domains, AI remains limited because rhetorical nuance, *dzawq lughawī*, intertextual meaning, and spiritual interpretation require human scholarly judgment. Therefore, the contribution of this study lies not in claiming that AI replaces traditional Arabic pedagogy, but in demonstrating how AI can operate within Islamic educational traditions.

Accordingly, this study offers two main contributions. Theoretically, it proposes an integrative model of AI-augmented Arabic learning that positions AI as a cognitive accelerator within a teacher-led, value-anchored ecosystem. This model clarifies the conceptual relationship among technological affordances, Arabic linguistic content, learner cognition, teacher authority, institutional readiness, and Islamic educational values. Practically, the study provides guidance for guardrailed prompting, teacher capacity-building, ethical AI-use policies, and institutional strategies for balancing digital innovation with *adab*, *sanad*, and authentic Arabic learning. By focusing on these two contributions, this article seeks to advance both global AI-education discourse and culturally grounded Arabic pedagogy in Southeast Asian Islamic institutions.

## **Method**

### **Research Design**

This study employed a convergent parallel mixed-methods design to examine AI integration in Arabic language learning within Islamic educational institutions. In this design, quantitative and qualitative data were collected during the same research phase, analyzed separately, and then integrated during interpretation to compare measurable patterns with contextual explanations (Creswell, 2021). This approach was appropriate because AI-mediated Arabic learning involves both observable learning outcomes and socio-pedagogical experiences shaped by institutional culture, teacher mediation, and learner engagement.

### **Research Context and Participants**

The study was conducted in five sites: *pesantren* in Wonosobo, *pesantren* in Bangil, UIN Antasari Banjarmasin, UII Darullughah Wadda'wah Pasuruan, and an Islamic educational institution in Kedah, Malaysia. These sites were selected purposively because they represent diverse Islamic Arabic learning environments in Indonesia and Malaysia, including *pesantren*, Islamic higher education, and traditional Malaysian Islamic education. The selection also

allowed the study to compare different levels of AI adoption, institutional readiness, and pedagogical orientation.

Participants were recruited using purposive sampling. The inclusion criteria were: students or santri actively involved in Arabic learning, teachers or lecturers teaching Arabic, prior exposure to AI-assisted learning tools, and willingness to participate voluntarily. Participants who were not involved in Arabic learning or did not provide consent were excluded. A total of 162 participants joined the study: Wonosobo pesantren (n = 34), Bangil pesantren (n = 31), UIN Antasari Banjarmasin (n = 36), UII Darullughah Wadda'wah Pasuruan (n = 29), and Kedah, Malaysia (n = 32). The participants included both male and female learners and instructors, with Arabic proficiency levels ranging from basic to upper-intermediate based on institutional placement and teacher evaluation.

### **Data Collection Techniques**

Data were collected through classroom observations, semi-structured interviews, focus group discussions, surveys, learning documents, and AI-generated learning artifacts. The AI tools observed included generative AI platforms, translation tools, grammar-support applications, speech-recognition tools, and Arabic NLP-based resources used for vocabulary, grammar, pronunciation, writing, and reading activities. This combination of data sources enabled triangulation between learner performance, classroom practices, teacher perceptions, and institutional responses.

### **Data Analysis Techniques**

Qualitative data were analyzed thematically by identifying patterns related to AI use, teacher mediation, learner autonomy, ethical boundaries, over-reliance, digital readiness, and the interpretation of classical Arabic texts. The analysis was informed by socio-technical systems theory, which views technology use as inseparable from institutional structures and human practices (Baxter & Sommerville, 2011), and by TPACK, which explains the interaction between technology, pedagogy, and content knowledge in teaching (Mishra & Koehler, 2006). Quantitative data from surveys and Arabic competency assessments were analyzed using descriptive statistics, including frequency, percentage, mean scores, and pre-post comparison. The qualitative and quantitative findings were then compared and integrated to explain how AI supported or constrained Arabic learning across different institutional contexts.

To ensure validity and trustworthiness, the study used methodological triangulation, cross-site comparison, member checking during FGDs, and comparison between survey findings, interview data, observation notes, and learning artifacts. Ethical procedures were followed by obtaining participant consent, anonymizing identities, and using AI-interaction records only for research purposes.

## Result and Discussion

This study generated findings through two complementary approaches: quantitative analysis via surveys and pre–post assessment of Arabic language competencies, and qualitative analysis through interviews, classroom observations, learning documents, and focus group discussions (FGDs) with lecturers, santri, and students from several Islamic educational institutions in Indonesia and Malaysia (particularly Kedah). Findings are presented across six major themes, revealing both advances and persistent challenges.

### Linguistic Performance Gains Through AI-Assisted Support

Quantitative analysis indicated measurable improvements in participants' linguistic performance following AI integration. Pre–post assessments covered vocabulary acquisition, syntax accuracy, reading comprehension, and pronunciation through AI platforms and teacher rubrics.

**Table 1. Summary of Arabic Language Competency Improvements Before and After AI-Assisted Learning**

| Competency              | Pre-test Mean | Post-test Mean | Mean Difference | Gain | Paired t-test   | p-value    | Effect Size |
|-------------------------|---------------|----------------|-----------------|------|-----------------|------------|-------------|
| Vocabulary              | 62.4          | 84.2           | 21.8            | +35% | $t(161) = 9.68$ | $p < .001$ | $d = 0.76$  |
| Nahwu & Sharaf Accuracy | 58.9          | 82.1           | 23.2            | +39% | $t(161) = 9.54$ | $p < .001$ | $d = 0.75$  |
| Reading Comprehension   | 65.7          | 86.8           | 21.1            | +32% | $t(161) = 9.16$ | $p < .001$ | $d = 0.72$  |
| Pronunciation Fluency   | 54.3          | 71.2           | 16.9            | +31% | $t(161) = 6.87$ | $p < .001$ | $d = 0.54$  |

Note. Paired-sample t-tests were used to compare pre-test and post-test scores. Effect size was calculated using Cohen's  $d$  for paired samples. Statistical significance was set at  $p < .05$ . Effect sizes indicate moderate to large improvements across the measured Arabic language competencies.

Quantitative analysis indicated statistically significant improvements in all measured Arabic language competencies after AI-assisted learning was introduced. Vocabulary scores increased from 62.4 to 84.2,  $t(161) = 9.68$ ,  $p < .001$ ,  $d = 0.76$ . Nahwu and sharaf accuracy improved from 58.9 to 82.1,  $t(161) = 9.54$ ,  $p < .001$ ,  $d = 0.75$ . Reading comprehension also

increased from 65.7 to 86.8,  $t(161) = 9.16$ ,  $p < .001$ ,  $d = 0.72$ . Pronunciation fluency improved from 54.3 to 71.2,  $t(161) = 6.87$ ,  $p < .001$ ,  $d = 0.54$ .

These results suggest that AI-assisted learning was associated with meaningful improvements in vocabulary acquisition, morphosyntactic accuracy, reading comprehension, and pronunciation fluency. However, the effect sizes indicate that the gains were moderate to large rather than automatic or uniformly strong. This finding supports the interpretation that AI can function as a useful micro-scaffolding tool for practice, feedback, and correction, but its impact remains dependent on teacher guidance, task design, and learners' ability to critically evaluate AI-generated output.

These data suggest that AI use contributed to accelerated linguistic understanding, particularly in morphological and syntactic pattern recognition, vocabulary expansion, voice-model-based pronunciation practice, and interactive text engagement. While these results align with findings from Zou et al. (2023) and Shi and Aryadoust (2024), our study underscores a critical nuance: improvement did not occur automatically but depended significantly on teacher guidance and goal-oriented learning strategies.

However, the linearity of these gains should not be overstated. Several instructors cautioned that students often appeared "more productive than they actually were," relying heavily on AI-generated structures without fully internalizing underlying grammar principles. This phenomenon what we term surface fluency resulted in students who could produce correct sentences yet struggled to explain their grammatical reasoning independently.

### **Self-Regulated Learning and the Paradox of Autonomy**

Survey responses indicated that 78% of participants felt more confident studying independently after using AI tools. Additionally: 74% used AI for self-directed vocabulary drills, 69% for sentence structure verification, 61% for conversational practice

Yet interview data revealed a more complex reality. One participant articulated:

**S-14**, Santri (IDN): *"AI makes us confident to study alone, but without guidance, sometimes we just follow the answers without truly understanding."*

This observation points to what we identify as *pseudo-competence* learners appearing proficient while lacking conceptual stability. The phenomenon parallels theories of cognitive dependency in AI-mediated environments (X. Wang et al., 2024). Students reported feeling empowered by immediate AI feedback, yet several teachers noted instances of "silent confusion": learners seemed active and engaged but could not articulate independent reasoning when AI prompts were unavailable.

The cultural dimension adds further nuance. In pesantren settings, several santri began treating AI as a *murāfiq ta'allum* (learning companion) rather than merely a tool, reflecting how Islamic learning identity adapts to digital mediation. However, educators at Malaysian

institutions stressed the necessity of ethical framing to prevent over-reliance and maintain *ikhhlās* (sincerity) in knowledge-seeking, echoing broader concerns in human-centered AI research (Alfredo et al., 2024).

Interpretation: AI accelerates certain forms of autonomy, yet without spiritual-ethical scaffolding and teacher oversight, this autonomy risks becoming hollow productive in appearance but shallow in substance.

### **AI's Limitations with Literary Language and Classical Texts**

A critical finding concerns AI's struggle with *balāghah* (rhetoric) and classical Arabic texts. While AI performed adequately with harakat (diacritical marks) in straightforward sentences, accuracy deteriorated markedly when encountering elevated literary language, poetry, and classical stylistics.

P-09, Student (IDN): "*AI can translate Arabic texts and vice versa, but for literary language and poetry, it is often inaccurate.*"

P-14, Teacher (IDN): "*For harakat placement it's often good, but when the language is too elevated and full of literary devices, it frequently errs. We cannot rely entirely on AI.*"

Specific weaknesses included: Misinterpreting *majāzī* (metaphorical) and *kināyah* (metonymic) meanings, failing to capture poetic rhythm and meter, missing spiritual undertones in religious texts, Lacking *dzawq lughawī* (linguistic taste) and aesthetic sensitivity

These limitations align with existing scholarship warning against full automation in classical Arabic interpretation (Alnajjar & Hämäläinen, 2024; Grigoryan et al., 2025). AI's algorithmic logic often cannot capture aesthetic inference, *maqāṣid al-lughah* (linguistic purposes), and semantic elegance that traditionally require trained human sensibility and spiritual awareness.

From a pesantren epistemological perspective, this limitation carries profound implications. Arabic language study is not merely instrumental but serves as a pathway to understanding revelation, ethics, and meaning purity. Thus AI's position becomes that of *khādim* (servant of knowledge), not *marja'* (authoritative reference). Classical texts (*turāth*) demand not only computational analysis but *tadabbur* (contemplative reflection) through *adab* (proper conduct), *sanad* (scholarly lineage), and teacher mentorship.

AI functions as a linguistic assistant but cannot serve as an authoritative judge in advanced Arabic or religious textual domains. Human scholarly intervention remains indispensable, particularly where authenticity, *adab*, and textual sanctity constitute core learning values.

## **Navigating Purpose and Boundaries: The Problem of Aimless Exploration**

A recurrent concern across research sites involved the risk of goal diffusion when students engaged with generative AI without clear task framing.

P-07, Lecturer (IDN): *"With ChatGPT, the advantages are significant, but we end up getting 'lost' unclear about goals and boundaries. There must be explicit guidelines on what information to seek and where the limits are."*

This observation encapsulates a broader pattern of *information drift*: students readily followed AI-suggested tangents, expanding inquiry beyond curricular aims without proportional learning depth. Teachers reported increased time-on-tool without corresponding gains when prompts lacked pedagogical constraints.

We interpret this as a design challenge rather than tool failure. Generative AI affords divergent thinking; absent defined purpose, retrieval scope, and stopping rules, such divergence can dilute depth. The concern also reflects Islamic educational ethics (*adab al-'ilm*): learning should be intentional (*niyyah*) and disciplined, not merely expansive.

Pedagogical Implication: Institutions should adopt "guardrailed prompting" frameworks: (1) state learning objectives and evidence criteria explicitly; (2) bound information sources (e.g., peer-reviewed, classical Arabic texts, authoritative fiqh references); (3) establish halt conditions (e.g., "compare three explanations then consult textbook"); and (4) require reflective closure linking AI outputs to course outcomes. Such scaffolds transform open-ended exploration into purposeful inquiry.

## **Traditional Learning Preferences in Kedah, Malaysia**

AI adoption in Kedah, Malaysia, exhibited notably different patterns. Implementation remained selective and functional rather than central to pedagogy.

P-18, Educator (MY-Kedah): *"In Kedah, understanding of AI is not yet comprehensive. We use it only as needed, according to our capacity. Traditional teaching methods remain our choice because we believe they are more assured and suit the madrasah atmosphere."*

This reflects what we term *prudential traditionalism* not resistance but reasoned caution grounded in pedagogical and spiritual values.

**Table 2. Contextual Comparison: Indonesia–Malaysia**

| <b>Dimension</b>        | <b>Indonesia (Pesantren)</b>       | <b>Malaysia (Kedah)</b>                  |
|-------------------------|------------------------------------|------------------------------------------|
| AI Adoption Pattern     | Progressive, integrative           | Selective, gradual                       |
| Teacher Role            | Central spiritual authority        | Central, cautious gatekeeper             |
| Pedagogical Orientation | Exploratory with pesantren filters | Traditional with limited experimentation |

|                  |                                    |                                                |
|------------------|------------------------------------|------------------------------------------------|
| Primary Concerns | Dependency, erosion of <i>adab</i> | Meaning distortion, loss of <i>rūḥ al-ʿilm</i> |
|------------------|------------------------------------|------------------------------------------------|

---

Three factors emerged as influential in Kedah's approach: Preference for *turāth* and *talaqqī*: Face-to-face learning with teachers carries perceived *barakah* (blessing), relational depth, and disciplinary rigor. Epistemic security concerns: Anxiety about AI errors in religious texts and elevated language. Technological capacity: Uneven AI literacy and developing infrastructure.

This "use-as-needed" approach does not signal rejection but rather maturity of judgment viewing AI not as replacement but as a tool that must submit to *adab*, *sanad*, and authentic learning needs. Such findings resonate with research indicating that technology adoption in tradition-based institutions often proceeds gradually, prioritizing cultural values, teacher authority, and spiritual dimensions of education (Alasmari, 2025; Sahrir et al., 2025).

AI integration in Kedah demonstrates that technology is adopted within frameworks of careful discernment grounded in religious values, maintaining harmony between digital innovation and traditional educational heritage.

**Ethical, Spiritual, and Epistemological Dimensions**

Participants expressed awareness of spiritual boundaries:

S-22, Santri (IDN): "*After using AI for a long time, we feel clever but empty. We must still read the classical texts (kitab).*"

P-05, Kyai (IDN): "*Machines help, but the light of knowledge comes from teachers. AI should not become our qiblah (direction of orientation).*"

These testimonies affirm that within Islamic educational frameworks, AI serves only as *khidmah* (service), not as an authority over meaning or scholarly lineage. Knowledge in this tradition involves not merely information transfer but *al-fahm* (deep understanding), *al-basīrah* (spiritual insight), and *al-tawqīr li al-ʿilm* (reverence for knowledge)—dimensions algorithmic systems cannot replicate.

From an Islamic epistemological perspective, this study verifies that AI functions safely only within frameworks where *ʿaql* (reason) remains directed and *naql* (transmitted knowledge) remains protected. High algorithmic accuracy does not replace the triad of understanding, insight, and reverence essential to authentic learning.

**Access Disparities and Digital Literacy Gaps**

Despite positive outcomes, structural challenges persisted:

**Table 3. Technical and Capacity Challenges**

| Challenge                              | Percentage |
|----------------------------------------|------------|
| Unstable internet connectivity         | 26%        |
| Uneven teacher digital literacy        | 22%        |
| Concerns about plagiarism & dependency | 18%        |
| Lack of clear AI usage guidelines      | 34%        |

Approximately 34% of respondents reported technical difficulties, while 27% expressed uncertainty in evaluating AI output correctness. These findings support global research indicating that AI success depends on institutional capacity, not technology alone (Chen et al., 2020; Hooda et al., 2022; B. Wang et al., 2023).

AI amplifies existing educational inequalities when adoption lacks scaffolded training and institutional support.

#### **Conceptual Synthesis: AI as Augmentor Within Tradition**

Taken together, findings affirm a balanced paradigm where AI and human teachers perform complementary yet distinct functions:

**Table 4. Complementary Roles of AI and Human Teachers in Islamic Arabic Language Learning**

| AI Provides              | Human Teachers Provide                           |
|--------------------------|--------------------------------------------------|
| Instant grammar feedback | Spiritual guidance & ethical grounding           |
| Vocabulary expansion     | Cultural context & <i>adab</i>                   |
| Adaptive drills          | Transformational mentorship                      |
| Simulated dialogues      | Authentic emotional & communal bonds             |
| Technical acceleration   | Meaning verification & <i>sanad</i> preservation |

Thus, AI operates as a cognitive amplifier and practice accelerator, while Islamic educators ensure epistemic integrity and embodied learning ethos—consistent with traditions emphasizing *ta'dib* (moral cultivation), *tarbiyah* (holistic education), and *ta'lim* (instruction).

The results affirm that AI transforms Arabic learning most effectively when combined with pesantren values, mentorship structures, and ethical formation. The breakthrough is not purely technological but pedagogical, cultural, and spiritual. AI accelerates skill acquisition, yet cultural-spiritual scaffolding remains vital. Human teachers remain central, not despite AI, but precisely because of it.

The quantitative results should not be interpreted merely as evidence of technical improvement, but as an indication of how AI functions within a broader socio-technical learning

ecology. The moderate to large effect sizes across vocabulary, nahwu–sharaf accuracy, reading comprehension, and pronunciation fluency suggest that AI can support Arabic learning when it provides immediate feedback, repeated practice, and individualized correction. From a TPACK perspective, these findings show that the pedagogical value of AI emerges when technological affordances are aligned with Arabic content knowledge and instructional design. AI was most useful not as an independent source of authority, but as a micro-scaffolding mechanism that helped learners notice errors, revise linguistic forms, and practice more frequently under teacher-guided learning objectives.

Overall, the findings indicate that AI should be positioned not merely as a technical tool, but as part of a socio-technical Arabic learning ecosystem. In line with socio-technical systems theory, the value of AI depends on the interaction between technology, human actors, institutional culture, and ethical governance (Baxter & Sommerville, 2011). The quantitative gains in vocabulary, nahwu–sharaf accuracy, reading comprehension, and pronunciation fluency show that AI can function as a cognitive amplifier through immediate feedback, repeated practice, and individualized correction. These findings support previous studies on AI-based feedback, speech evaluation, and Arabic writing support (Shi & Aryadoust, 2024; Zou et al., 2023; Zubaidi et al., 2025), while also extending them by showing that, in Islamic Arabic education, AI gains become meaningful only when guided by teachers and institutional values.

From a TPACK perspective, AI becomes pedagogically useful when technological affordances are aligned with Arabic content knowledge and instructional design (Mishra & Koehler, 2006; Priyanda et al., 2025). This aligns with broader AI education studies showing that generative AI can support learning, writing, feedback, and accessibility, but also raises concerns about accuracy, dependency, academic integrity, and unequal access (Chan & Hu, 2023; Chiu et al., 2023; Crompton & Burke, 2023; Farrokhnia et al., 2024; Kasneci et al., 2023; Lo, 2023; Ouyang & Jiao, 2021; Su & Yang, 2023; Tlili et al., 2023). In this study, AI was most effective when used for structured micro-scaffolding, such as vocabulary drills, sentence correction, pronunciation rehearsal, and guided reading. Without clear objectives, however, it could produce goal diffusion, over-reliance, and surface fluency.

The findings also contribute to discussions on learner agency and psycholinguistic processing. AI feedback may help learners notice errors and revise linguistic forms, but feedback becomes meaningful only when learners internalize rules and can explain their reasoning beyond AI-generated answers (Levelt, 2012; X. Wang et al., 2024). This supports studies showing that AI assistance can strengthen student agency while also creating risks of passive dependency if not pedagogically regulated (Alfredo et al., 2024; Darvishi et al., 2024).

Therefore, AI-supported autonomy in Arabic learning must be balanced with teacher mediation, reflective explanation, and *adab al-'ilm*.

The contextual findings show that AI adoption is shaped by sociolinguistic and institutional realities. Arabic learning in *pesantren*, Islamic universities, and Malaysian Islamic institutions is embedded in community norms, teacher authority, and identity formation (Wardhaugh & Fuller, 2014). This supports recent studies showing that Arabic learning in Southeast Asia is strongly influenced by immersive environments, informal sociolinguistic interaction, motivation, and structured technology use (Kasmaluddin et al., 2026; Muhith & Batrisya, 2026; Mukhasibi & Wahyudi, 2026; Sahrir et al., 2025). Thus, AI may expand practice opportunities, but it cannot replace the relational, ethical, and communal dimensions of Arabic learning.

Practically, these findings suggest that Islamic institutions should institutionalize AI use through guardrailed prompting, teacher AI-literacy training, ethical AI policies, and reflective learning tasks. AI literacy is necessary because teachers and learners must be able to evaluate AI outputs, identify inaccuracies, and avoid dependency. Wang et al. (2023) identify AI literacy as a key competence in human–AI interaction, while Kohnke et al. (2023) emphasize that ChatGPT use in language learning requires pedagogical awareness and critical evaluation.

Responsible AI adoption also requires institutional governance. Miao and Holmes (2023) stress the need for human-centered and ethical principles in generative AI use, while Hooda et al. (2022) argue that AI-based feedback is effective only when embedded in responsible learning systems. Hwang and Chang (2023) similarly note that educational chatbots offer interaction and feedback, but require careful risk management.

In Arabic learning, students should verify AI outputs through textbooks, teacher explanation, peer discussion, and classical references. Godwin-Jones (2022, 2023) argues that AI-based language tools can expand practice opportunities, but still require teacher mediation and critical digital literacy. Liang et al. (2023) show that AI is widely used for feedback, writing support, and adaptive language learning, while Xiao et al. (2025) emphasize the need for structured pedagogical design.

AI implementation must consider acceptance, equity, and sustainability. Strzelecki (2024) shows that ChatGPT acceptance is shaped by usefulness, ease of use, and trust, whereas Weng and Fu (2025) emphasize that generative AI should bridge learning divides rather than widen them. Therefore, Islamic institutions need teacher training, student guidance, ethical regulation, and culturally appropriate AI-use models aligned with *adab*, *sanad*, teacher authority, and epistemic integrity.

In this study, AI integration refers to the pedagogically guided use of generative AI, translation tools, grammar-support applications, speech-recognition tools, and Arabic NLP-

based resources to support Arabic learning tasks. Micro-scaffolding refers to small, targeted forms of AI assistance, such as error detection, vocabulary recall, sentence correction, pronunciation rehearsal, and guided reading support. A value-anchored learning ecosystem refers to an instructional environment in which AI use is regulated by teacher mediation, institutional ethics, and Islamic educational values such as *adab*, *sanad*, *niyyah*, and epistemic integrity. These conceptual definitions clarify that AI is not treated as an autonomous instructional authority, but as a supporting mechanism within a teacher-led Arabic learning system.

Finally, the limitations of AI in *turāth*, *balāghah*, poetry, and religious discourse confirm the continued importance of human scholarly judgment. Although Arabic NLP and speech technologies continue to develop (Alnajjar & Hämäläinen, 2024; Grigoryan et al., 2025), AI still struggles with *majāz*, *kināyah*, intertextual meaning, rhetorical nuance, and *dzawq lughawī*. Therefore, AI should be understood as *khādim al-ta'allum*, a servant of learning, not *marja' al-ma'nā*, an authority over meaning. The main contribution of this study is to show that AI in Islamic Arabic education is most valuable as disciplined augmentation: technology accelerates practice, while teachers and Islamic traditions preserve the depth, direction, and dignity of knowledge.

## **Conclusion**

This study examined how artificial intelligence (AI) transforms Arabic language learning in Islamic educational institutions across Indonesia and Malaysia through a convergent parallel mixed-methods design. The findings show that AI can support vocabulary acquisition, *nahwu-sharaf* accuracy, reading comprehension, and pronunciation fluency when used as a micro-scaffolding tool for feedback, practice, correction, and guided interaction. However, these benefits are not automatic. AI becomes educationally meaningful only when embedded within teacher mediation, clear learning objectives, ethical boundaries, and institutional values. The theoretical novelty of this study lies in its integrative socio-technical model, which positions AI not as an autonomous pedagogical authority but as a cognitive accelerator within a teacher-led, value-anchored Arabic learning ecosystem. By connecting TPACK, socio-technical design, psycholinguistic processing, sociolinguistic context, and Islamic educational values such as *adab*, *sanad*, *niyyah*, and epistemic integrity, this study contributes a conceptual framework for understanding AI integration in Arabic education beyond technical effectiveness.

Contextually, this study contributes to Southeast Asian Islamic education by showing that AI adoption is shaped by institutional identity, religious learning traditions, teacher authority, and cultural-pedagogical readiness. Indonesian sites tended to integrate AI more progressively through *pesantren*-based filters, whereas the Kedah context reflected a more

selective and cautious model rooted in talaqqī, turāth learning, and epistemic security. Practically, the study recommends guardrailed prompting, teacher AI-literacy development, ethical AI-use policies, reflective learning tasks, and clear role division between AI and human teachers. Its limitations include reliance on descriptive and moderate inferential pre–post data, possible variation in assessment rubrics across sites, and limited fine-grained analysis of dialectal, rhetorical, and prosodic dimensions of Arabic. Future research should examine longitudinal learning transfer, dialect-sensitive and balāghah-aware NLP models, oral defense tasks to reduce pseudo-competence, and scalable institutional policies for equitable AI integration. Looking ahead, the future of Arabic education in Islamic institutions should not be framed as AI versus tradition, but as AI with tradition: a model in which intelligent technologies accelerate practice, while teachers, texts, and Islamic scholarly values preserve the depth, dignity, and direction of knowledge.

## References

- Alasmari, T. (2025). Artificial Intelligence and M-Learning in Arabic Countries: Innovations, Trends, and Regional Perspectives. *International Journal of Interactive Mobile Technologies (IJIM)*, 19(05), 170–194. <https://doi.org/10.3991/ijim.v19i05.52735>
- Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 6, 100215. <https://doi.org/10.1016/j.caeai.2024.100215>
- Alkaabi, M. H., & Almaamari, A. S. (2025). Generative AI Implementation and Assessment in Arabic Language Teaching: *International Journal of Online Pedagogy and Course Design*, 15(1), 1–18. <https://doi.org/10.4018/IJOPCD.368037>
- Alnajjar, K., & Härmäläinen, M. (2024). Normalization of Arabic Dialects into Modern Standard Arabic using BERT and GPT-2. *Journal of Data Mining & Digital Humanities, NLP4DH*, 13146. <https://doi.org/10.46298/jdmdh.13146>
- Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, 23(1), 4–17. <https://doi.org/10.1016/j.intcom.2010.07.003>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Creswell, J. W. (2021). *A Concise Introduction to Mixed Methods Research*. SAGE Publications, Inc.
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the

- field. *International Journal of Educational Technology in Higher Education*, 20(1), 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, 210, 104967. <https://doi.org/10.1016/j.compedu.2023.104967>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474. <https://doi.org/10.1080/14703297.2023.2195846>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent Writing Assistance and Instructed Language Learning. *Language Learning & Technology*, 26(2), 5–24.
- Godwin-Jones, R. (2023). Emerging Spaces for Language Learning: AI Bots, Ambient Intelligence, and the Metaverse. *Language Learning & Technology*, 27(2), 6–27.
- Grigoryan, L., Karpov, N., Albasiri, E., Lavrukhin, V., & Ginsburg, B. (2025). Open Automatic Speech Recognition Models for Classical and Modern Standard Arabic. *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. <https://doi.org/10.1109/ICASSP49660.2025.10889643>
- Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education. *Mathematical Problems in Engineering*, 2022, 1–19. <https://doi.org/10.1155/2022/5215722>
- Hwang, G.-J., & Chang, C.-Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 31(7), 4099–4112. <https://doi.org/10.1080/10494820.2021.1952615>
- Kasmaluddin. (2026). Duolingo-Based Arabic Vocabulary Acquisition in Islamic Educational Contexts. *Al-Muhawwaroh: Jurnal Pendidikan Bahasa Arab*, 2(1), 1–11. <https://doi.org/10.38073/almuhawwaroh.v2i1.4358>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Levelt, W. (2012). *A History of Psycholinguistics: The Pre-Chomskyan Era*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199653669.001.0001>
- Liang, J.-C., Hwang, G.-J., Chen, M.-R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*, 31(7), 4270–4296. <https://doi.org/10.1080/10494820.2021.1958348>
- Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- Miao, F., & Holmes, W. (2023). *Guidance for Generative AI in Education and Research*. UNESCO.

- Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record: The Voice of Scholarship in Education*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Muhith, M., & Batrisya, N. A. (2026). ‘Lepak’ as a Sociolinguistic Arena: Oral Interaction and Arabic L2 Fluency in Northern Malaysia. *Farasyah: Journal of Linguistics and Language Education*, 1(1), 19–29.
- Mukhasibi, A., & Wahyudi, I. (2026). Immersive Arabic Islamic Environment and Student Engagement: Evaluating the Linguistic Impact of World Arabic Language. *Farasyah: Journal of Linguistics and Language Education*, 1(1), 1–12.
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020. <https://doi.org/10.1016/j.caeai.2021.100020>
- Priyanda, R., Herman, T., Amalia, R., & Ihsan, I. R. (2025). Exploring teachers' pedagogical reasoning in mathematics education using the TPACK framework. *Frontiers in Education*, 10, 1552760. <https://doi.org/10.3389/feduc.2025.1552760>
- Sahrir, M. S., Muharom Albantani, A., Arifin, F., & Din, F. (2025). The Use of Generative Artificial Intelligence (AI) in Arabic Language Education: Insights and Implications between Malaysia and Indonesia. *International Journal of Research and Innovation in Social Science*, IX(III), 3638–3646. <https://doi.org/10.47772/IJRIS.2025.90300288>
- Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, 36(2), 187–209. <https://doi.org/10.1017/S0958344023000265>
- Strzelecki, A. (2024). Students' Acceptance of ChatGPT in Higher Education: An Extended Unified Theory of Acceptance and Use of Technology. *Innovative Higher Education*, 49(2), 223–245. <https://doi.org/10.1007/s10755-023-09686-1>
- Su, J., & Yang, W. (2023). Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education. *ECNU Review of Education*, 6(3), 355–366. <https://doi.org/10.1177/20965311231168423>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, 42(9), 1324–1337. <https://doi.org/10.1080/0144929X.2022.2072768>
- Wang, X., Xu, X., Zhang, Y., Hao, S., & Jie, W. (2024). Exploring the impact of artificial intelligence application in personalized learning environments: Thematic analysis of undergraduates' perceptions in China. *Humanities and Social Sciences Communications*, 11(1), 1644. <https://doi.org/10.1057/s41599-024-04168-x>
- Wardhaugh, R., & Fuller, J. M. (2014). *An Introduction to Sociolinguistics*. Wiley.
- Weng, Z., & Fu, Y. (2025). Generative AI in Language Education: Bridging Divide and Fostering Inclusivity. *International Journal of Technology in Education*, 8(2), 395–420. <https://doi.org/10.46328/ijte.1056>
- Xiao, Y. (2025). The impact of AI-driven speech recognition on EFL listening comprehension, flow experience, and anxiety: A randomized controlled trial. *Humanities and Social Sciences Communications*, 12(1), 425. <https://doi.org/10.1057/s41599-025-04672-8>

- Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An Investigation Into Artificial Intelligence Speech Evaluation Programs With Automatic Feedback for Developing EFL Learners' Speaking Skills. *Sage Open*, 13(3), 21582440231193818. <https://doi.org/10.1177/21582440231193818>
- Zubaidi, A., Munip, A., Widodo, S. A., & Zerrouki, T. (2025). Enhancing Arabic writing skills using Chat GPT-based AI learning models: A tridimensional human-AI collaboration framework. *Indonesian Journal of Applied Linguistics*, 15(1), 87–101. <https://doi.org/10.17509/ijal.v15i1.75378>