# artikel ijaz (1).docx

*by* User User

**Submission date:** 31-Mar-2024 04:58AM (UTC+0300)
**Submission ID:** 2335559678
**File name:** artikel_ijaz_1_.docx (186.58K)
**Word count:** 3561
**Character count:** 21654

# AI Concepts Integration in Developing E-Muhadathat Kits for Non-Arabic Speakers: An Overview

**Abstract**

Contemporary educational reforms emphasise the strategic integration of technology into pedagogy and are in line with the principles of the fourth industrial revolution (IR 4.0). Among the rapidly advancing technologies in the field of language learning is Artificial Intelligence (AI), also known as Natural Language Processing (NLP). This technology has been incorporated in the development of the E-Muhadathat kit, an interactive Arabic conversation simulation model that aims to improve the communication skills of non-Arabic speaking students in Malaysian public universities. The aim of this study is to explore the concept of AI in education, address the challenges in developing AI-based Arabic language software and propose a model for the E-Muhadathat kit that is customised for non-Arabic speakers. This literature review uses documentary methods to collect information from journals, conference proceedings, articles, dissertations and digital books retrieved from databases such as Google Scholar, Springer Link, Science Direct and Research Gate. The collected data is then subjected to a descriptive analysis based on thematic categorisation. The results show that the concept of AI comprises three main components: machine learning, deep learning and neural networks. Arabic linguists have identified various challenges in the development of AI-based software, including orthographic, morphological, syntactic, anaphoric and semantic ambiguities. In addition, the design of the E-Muhadathat kit incorporates elements of machine learning and deep learning and includes applications for both spoken and written language. The findings of this study are of significant help to educators and researchers who intend to develop Arabic language software that focuses on the communication of non-native speakers in Malaysian higher education institutions. Consequently, the study recommends further exploration of AI in Arabic language teaching and learning to improve the communication skills of non-native speakers in Malaysia.

**Keywords:** Artificial Intelligence; Natural Language Processing; E-Muhadathat; Arabic teaching and learning; non-Arabic speakers

## INTRODUCTION

In today's sophisticated digital landscape, Artificial Intelligence (AI) technology enables machines to observe and learn from human behavior. This technology utilizes software programs based on algorithms or robots that can understand, interpret, and respond to human commands. Its extensive utilization has resulted in the creation of prevalent high-tech gadgets like smartphones, tablets, laptops, autonomous vehicles, drones, and robots (Nawi, 2019). Furthermore, popular products in the market, such as Amazon Alexa, Siri-Apple, OK Google, or Hey Google applications, are employed to recognize human voices for tasks such as playing music, locating tourist destinations, and controlling home devices (Naik, 2020). AI technology, including its subset Natural Language Processing (NLP), finds widespread application in language learning. Notable examples include Google Translate for multilingual translation, Chatbot and Wit.ai for interactive conversations in both spoken and written forms, Gmail's auto-mail feature, and Grammarly for automated text prediction and sentence correction. In the realm of

Arabic language pedagogy, the utilization of this technology is still in its early stages and warrants further exploration (Khalati & Al-Romany, 2020; Luthfiah, 2023). Several factors contribute to this, including a scarcity of scholars or experts in Arabic language computing, the diversity of Arabic language dialects, and a delayed awareness of leveraging internet networks and machine-assisted translation to bolster Arabic language proficiency (Anwar & Ahyarudin, 2023; Adel, 2022). Researchers have given limited attention to the development of media kits for Arabic communication skills using AI technology (Abdelghani, 2021; Abdurrosid & Mamluatul, 2021). This article aims to fill this gap by exploring the intersection of AI and education. The main objective is to establish the foundation for creating the E-Muhadathat kit model, which is specifically designed to enhance Arabic communication skills for non-native speakers. This interactive conversation simulation model, which can be accessed via smartphones, helps non-native student's master Arabic communication. Furthermore, this article identifies and addresses the challenges involved in developing Arabic language software that integrates AI technology, particularly in creating the E-Muhadathat kit model. Additionally, it proposes a model design that utilizes the ADDIE Instructional Model procedure, which includes analysis, design, development, implementation, and evaluation. This tailored approach aims to benefit non-Arabic speakers within Malaysian Higher Education Institutions.

**PROBLEM STATEMENT**

According to Sahrir (2016), the importance of the Arabic language extends beyond the Muslim community to encompass various professional domains and international communication channels such as translation, business, diplomatic relations, broadcasting, and mass media. This widespread use of Arabic in the field of science and technology has a positive effect on producing professionals at the global level (Ramli et al., 2023; Romli et al., 2018; Atoh et al., 2014). On the other hand, Khalati and Al-Romany (2020) argue that measures to empower the Arabic language at the global level need to be expanded the efforts to empower the Arabic language globally need to be expanded, considering Arabic's status as the fifth most spoken language worldwide, with more than 400 million speakers (Cabral, 2023; Kaseh et al., 2023).

However, Arabic's presence on the internet remains limited, accounting for only about 1% of online content compared to its speakers, who make up 5% of the global population. As stated by Abdurrosid and Mamluatul (2021), the Arabic language should be utilized more broadly across different media platforms using NLP concepts, such as websites, machine-assisted translation, smartphone applications, and social media. This expansion aims to help speakers, including non-Arabic speakers, to master this international language, given its importance in professional and international communication fields. In Malaysia, there is a growing trend of integrating multimedia elements into Arabic teaching and learning to create a student-centered, interactive learning environment. This is attributed to the conventional approach, which predominantly relies on textbooks and tends to be teacher-centered in the teaching and learning process of the Arabic language (Ramli et al., 2019; Romli et al., 2018; Ramli, et al., 2017).

Traditional methods, relying mainly on textbooks and teacher-centered approaches, are now being supplemented with innovative techniques. By integrating educational technologies such as pictures, audio, video, text, and animation as teaching aids, student interest can be sparked, leading to enhanced Arabic speaking skills. This,

in turn, fosters more effective Arabic pedagogical activities (Ghani & Daud, 2023; Zakaria et al., 2019). Therefore, the development of an E-Muhadathat kit, designed as a smartphone application, can offer students valuable opportunities to partake in authentic conversations with Arabic speakers. Such interactions serve to deepen their proficiency in the language and refine their communication skills.

## METHOD

In this study, a literature review approach is adopted, focusing on library resources. The research methodology employed is qualitative and descriptive, entailing structured analysis of collected data presented in a scientific report format. The data sources encompass diverse library materials, such as books, journals, papers, and articles relevant to the topic under consideration. Data collection involved utilizing specific keywords, including "Artificial Intelligence," "Arabic Natural Language Processing (ANLP)," "Natural Language Processing (NLP)," "teaching and learning Arabic," "communication skills in Arabic," and "Arabic communication skills," to conduct searches across databases like Google Scholar, Springer Link, Science Direct, and Research Gate. Subsequently, the gathered online data underwent thematic categorization and descriptive analysis to derive insights.

## RESULTS AND DISCUSSION

### Concept of Artificial Intelligence (AI)

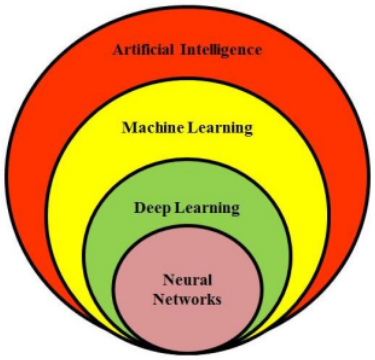
AI technology is a remarkable field that aims to replicate human intelligence, serving to solve various challenges and problems. Its three key techniques, Fuzzy Logic (FL), Evolutionary Computing (EC), and Machine Learning (ML), are transforming industries and changing our lives (Ahmad, 2017). FL techniques replicate human reasoning, considering all potential outcomes and interpreting between binary values of 0 and 1 (Nasution, 2012). The automatic braking systems employed in Japanese trains are a perfect example of FL application, ensuring passenger safety and preventing accidents. EC involves computational evolution, emphasizing natural evolutionary strategies to achieve optimal or near-optimal problem-solving. It is used in air conditioning compressor and inverter systems, enabling automatic adjustment of room temperature to conserve electricity. It is a game-changer in energy conservation and efficiency.

ML, the most widely used technique, is the core of AI advancements, enabling software to make more accurate predictions based on experience without explicit programming. It has numerous applications, from fraud detection, and spam filtering to malware threat detection and more. It is the backbone of many businesses, driving growth and innovation. Furthermore, AI technology is built on three fundamental concepts: machine learning, deep learning, and neural networks. Machine learning mimics human intelligence by learning from its surroundings. It is a powerful tool that has revolutionized various domains such as spacecraft engineering, medicine, pattern recognition, and more, by leveraging Big Data (El Naqa & Murphy, 2015).

On the other hand, deep learning enables computers to perform complex tasks naturally using images, text, and voice and is responsible for remarkable advancements such as driverless cars, image classification, voice recognition, machine-assisted translation, and many others (Nawi, 2019). These cutting-edge technologies are changing the world as we know it, and their potential is limitless. This concept has the potential to attain remarkable accuracy, sometimes even surpassing human intelligence.

Moreover, the concept of neural networks employs a reflective system inspired by the functioning of the human brain. This enables computer programs to recognize patterns and emulate the communication between biological neurons. Applications of neural networks include rain detection, currency authenticity verification, and disease detection via X-ray scans, depression identification, and more (Annisa et al., 2020). A summary of the three main concepts of AI is depicted in Figure 1 below:

Figure 1. Main Concepts of AI in Education




**Challenges and Obstacles in Developing Arabic AI Software**

The rapid development of Artificial Intelligence (AI) has brought computers to a level of capability that rivals that of humans. AI systems with Natural Language Processing (NLP) capabilities can now process natural human language on the internet, including expressions in Arabic. Despite this progress, the full potential of Arabic language usage on virtual platforms remains limited due to several issues and challenges in the development of AI software. Experts in Arabic Natural Language Processing have highlighted various challenges, including the use of Chatbot to aid teachers in the teaching and learning process of Arabic (Almurayh, 2021).

Recent studies indicate a growing interest in AI technology among both Arabic and non-Arabic speakers, particularly in the automatic recognition of threatening, cyber bullying, and hateful texts in Arabic on social media (Elzayady et al., 2022). However, challenges still exist in using algorithms to help Arabic students overcome morphological errors and ambiguities (Elayeb et al., 2022). ANLP scholars continue to focus on critical issues such as orthographic, morphological, syntactical, anaphoric, and semantic ambiguities (Shaalan et al., 2018).

***Orthographic Ambiguity***

Arabic Orthographic Ambiguity refers to uncertainty in orthographic symbols, commonly known as diacritics, which are marks or symbols used to differentiate sounds within a letter. In Arabic, diacritics serve to clarify the pronunciation of letters and the meanings of words, also known as tashkeel or short vowels, such as the use of lines in words like كَتَبَ (he has written), كُتِبَ (has been written), and كُتُبٌ (books). Ambiguous orthography can render Arabic text illegible and compromise the intended meaning of a sentence. However, this ambiguity can be mitigated by incorporating specific diacritics

into ANLP processing for text-to-speech or conversation (Al mekhlafi et al., 2022; Jarrar et al., 2018).

***Morphological Ambiguity***

Arabic is renowned for its exceptionally rich morphology, characterized by its unique template known as "Mizan Sorfi." The complexity of Arabic morphology often leads to ambiguity in distinguishing between word and sentence structures due to its high flexibility and intricacy. However, this complexity can be navigated through specific derivation processes tailored to different levels of proficiency. Arabic word derivation entails the utilization of three or four-character root words, depicted in various patterns and augmented with affixes. A concise overview of word derivation in Arabic is provided in Table 1 below:

**Table 1.** Illustration of Arabic Word Derivation Levels

| Words | Transliteration | Root Words | Transliteration | Pattern |
|---|---|---|---|---|
| حافظ | Hãfiz | حفظ | Hafaza | _ _ ا _ |
| كاتب | Kãtib | كتب | Kataba | _ _ ا _ |
| معلوم | Ma'lũm | علم | 'Alima | م _ _ و _ |
| مدروس | Madrũs | درس | Darasa | م _ _ و _ |

**Source:** Shaalan et al., 2018

***Syntax Ambiguity***

Arab linguists established the foundation of Arabic grammar rules to prevent errors in Quran recitation. However, these grammar rules pose challenges for computers in automatically processing Arabic text with ANLP. Verbal and nominal sentences in Arabic have a different structure than that in English. For example, in the verbal sentence "يتعلم التلميذ" (the student is studying; yata'allamu al-Talib), the verb comes before the subject, unlike in English where all sentences begin with a subject followed by a verb. Nominal sentences in Arabic begin with a noun or pronoun and can be classified into two types: those without verbs and those with verbs. Nominal sentences without verbs imply the concept of "to be," similar to the English "verb to be," as demonstrated by phrases like "the weather is wonderful." However, in Arabic, the verb is not explicitly used, as seen in the example "الجو جميل" (beautiful weather).

***Anaphoric Ambiguity***

Utilizing anaphora resolution in Arabic involves associating a particular entity or pronoun with the corresponding noun or name to which it refers. Examples of anaphora usage in Arabic sentences as shown in Table 2 below:

**Table 2.** Anaphora in Arabic Sentence

| Arabic Sentence | Machine Translation |
|---|---|
| رأيت القطة، فاعطيتها الطعام | I saw the cat, so I gave it food |
| رأيت القطة، فاعطيتها الطعام | I saw the cat, so I gave her food |

**Source:** Shaalan et al., 2018

Referring to Table 2, the comparison between the phrases 'give it' and 'give her' illustrates the ambiguity of anaphora in Arabic. In Malay or English, the appropriate term would be 'give it.' This discrepancy arises from the classification of the noun for cats as an animal, unlike the use of 'her' for nouns referring to people. Machine

translation encounters challenges in processing the anaphora for the third person pronoun or ها which refers to the noun cat in the aforementioned sentence.

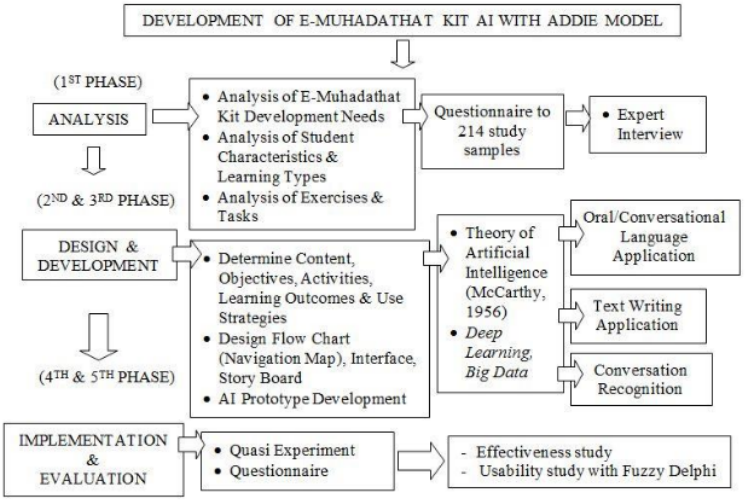
***Semantic Ambiguity***

The semantic ambiguity inherent in the Arabic language encompasses sentences or phrases that require contextual adaptation to convey the correct meaning. For instance, consider the verse "يحب علي أحمد أكثرمن إبراهيم" (Ali loves Ahmad more than Ibrahim). This raises the question: does it imply that Ali loves Ahmad more than Ibrahim does, or that Ali loves Ahmad more than he loves Ibrahim? Consequently, the semantic ambiguity of Arabic poses a significant challenge in developing ANLP Arabic language software capable of discerning sentence meanings across various contexts.

***Model Design of E-Muhadathat AI Kit***

The development of an E-Muhadathat kit, which is an AI-powered interactive Arabic conversation simulation model for smartphones, requires a well-structured framework that can serve as a reference for researchers. This study follows the ADDIE Instructional Model procedure (Branch, 2014), which delineates five key stages: analysis, design, development, implementation, and evaluation as Figure 2 below.

Figure 2. Overview of the E-Muhadathat AI Kit Model




Based on Figure 2, this study employs the ADDIE Model procedure in the development of E-Muhadathat kits for non-Arabic speakers. This iterative model ensures thorough and systematic development, with comprehensive evaluations conducted at each phase (Abu et al., 2020; Arasinah & Amarumi, 2021). The study encompasses five key stages aimed at developing the E-Muhadathat kits effectively. The first stage involves analysis. Here, the focus lies on analyzing the development needs of the E-Muhadathat kits, analyzing student characteristics and learning preferences, and assessing exercises and tasks related to Arabic communication and speaking skills. This phase incorporates a questionnaire administered to 214 study

participants and expert interviews to identify the specific requirements for kit development.

The 2nd and 3rd phase emphasize the design and development process which consist of determination of content, objectives, activities, learning outcome and strategies use in teaching and learning. It comprises the design of flow chart (navigation map), interface, story board and the development of AI kits prototype. This phase also adapts theory of artificial intelligence by McCarthy (1956) and incorporates key AI concepts such as machine learning and deep learning. It enables the processing of written text, spoken conversation, and text-to-speech functionality with a spoken conversation recognition system. Moreover, the 4th and 5th phase focuses on implementation and evaluation which uses usability test to experts with Fuzzy Delphi Method (FDM) and quasi experiment to non-Arabic students in public universities.

Referring to the literature review, it is evident that AI technology significantly contributes to the enhancement of the Arabic language presence online. This advancement facilitates the dissemination of Arabic content across the internet, catering to diverse age groups ranging from children to adults. Furthermore, the utility of the Arabic language extends beyond religious boundaries, finding relevance in professional spheres and international discourse (Sahrir, 2016). Addressing the challenges inherent in Arabic AI software development, particularly about linguistic ambiguities, such as orthographic, morphological, syntactic, anaphoric, and semantic ambiguities, is paramount. A notable solution lies in the integration of specific diacritics or tashkeel to clarify letter pronunciation and word meanings (Almekhlafi et.al, 2022).

This also supports Nazih and Hifny's (2022) study that has been developed Arabic BERT Models that represents a significant stride toward mitigating orthographic ambiguity in ANLP processing. These models provide a robust framework for enhancing the accuracy and efficacy of AI-driven language processing systems tailored for Arabic, thereby addressing crucial linguistic challenges in the digital realm (Obied et al., 2022). Furthermore, Arabic morphological ambiguity can be effectively addressed through a staged derivation process, as outlined in the Mizan Sorfi template.

This methodology aligns with the Camelira system to mitigate morphological ambiguity across various dialects (Elayeb et al., 2022). This morphological ambiguity can be solved by investigating strategies to tackle morphological ambiguity encompassing word combination, derived word, and text ambiguity, without resorting to symbols or lines. Their work culminated in the development of a machine learning algorithm tailored to aid non-Arabic speakers in navigating morphological ambiguity within hadith texts.

Moreover, syntactic ambiguity, including the complexities inherent in verbal and nominal sentences, poses another challenge. Arabic syntactic ambiguity can be resolved through the application of matching formulas for rationality, gender, and number between the inchoative or المبتدأ in nominal sentences. Similarly, in verbal sentences, syntactic ambiguity can be alleviated by employing verb-subject and verb-agent matching formulas (Othman et al., 2004). This supports the study of Almanna and Jamoussi (2022) who have developed the Neural Machine Translation (NMT) system to facilitate precise translations between English and Arabic sentences, thereby aiding in the comprehension and resolution of syntactic ambiguity.

One of the challenges of Arabic language processing is anaphora ambiguity, which can be overcome by matching entities' specific pronouns or referred names (Shaalan et al., 2018). To address this issue, a set of computational and linguistic

features have been developed that use machine learning algorithms (Abolohom et al., 2021). Another challenge is the semantic ambiguity of the Arabic language, which requires matching sentences or phrases with different contexts to produce the correct meaning. To overcome this, semantic marking can be used to relate text data elements to a well-formed ontology or lexicon (Rayson et.al, 2004).

A lexicon is a component of an NLP system that contains semantic or grammatical information about single words or sets of words. Thus, semantic ambiguity can also be overcome with semantic marking. This supports the study El-Haj et.al (2022) has developed Arabic semantic marking system (AraSAS) with Python 3 has been developed to convert raw data (Arabic text) to semantic marking, enabling the appropriateness of a sentence or phrase to fit its meaning in different contexts to be determined. Although developing Arabic language software with AI technology, especially ANLP, presents various challenges, these hurdles must be addressed through collaborative efforts.

It is crucial to engage experts in AI, computer-assisted language learning (CALL), innovation in language education technology, and other related fields to find effective solutions for Arabic software development based on AI technology. Furthermore, there is a critical need to improve the integration of AI concepts into the development of educational software, such as the E-Muhadathat development kit. Taking inspiration from research endeavors, we must further develop this exploration to encompass other Arabic language skills. These skills include Maharat al-Istima' (listening skills), Maharat al-Qira'ah (reading skills), and Maharat al-Kitabah (writing skills). By leveraging AI advancements in these areas, educational software can offer more comprehensive and effective learning experiences for Arabic language learners.

## CONCLUSION

In Malaysia, there is a need to incorporate more AI concepts into the development of software for teaching the Arabic language. The key concepts include machine learning, deep learning, and neural networks, which are crucial in creating educational software with AI technology, such as websites, robotics, smartphone applications, and other mobile devices. However, developing AI-based Arabic language software comes with its challenges, such as orthographic, morphological, syntactic, anaphora, and semantic ambiguity. Therefore, AI and Arabic language experts must investigate these issues thoroughly, particularly in the design and development phases of Arabic language education software. The study emphasizes the importance of the E-Muhadathat AI kit model, which applies machine learning and deep learning concepts to improve spoken language comprehension, written text analysis, and conversational recognition. This model can provide valuable insights into the development of Arabic language content modules and online learning models tailored for non-native speakers in Malaysian Higher Education Institutions (HEIs). Furthermore, by embracing student-centered interactive Arabic learning approaches, educators can better support learners in their language acquisition journey. Finally, the study advocates for the ongoing exploration and development of AI in Arabic language teaching and learning. This exploration is crucial in enabling students to communicate effectively with non-Arabic speakers in Malaysian universities. By embracing AI technologies, educators can create an enriched and inclusive learning environment conducive to cross-cultural communication and understanding.

# artikel ijaz (1).docx

ORIGINALITY REPORT

| 8% | 7% | 2% | 1% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | ojs.upsi.edu.my<br>Internet Source | 4% |
| 2 | purehost.bath.ac.uk<br>Internet Source | 1% |
| 3 | Alkhatib, Manar. "Neural Machine Translation for Arabic Language", The British University in Dubai, 2023<br>Publication | 1% |
| 4 | ejournal.radenintan.ac.id<br>Internet Source | 1% |
| 5 | Liuqiao Sun, Hangjun Chen, Qian Ou, Xiaoping Liang et al. "The effects of nutritional interventions in patients with non-alcoholic fatty liver disease: An umbrella review of meta-analyses of randomized controlled trials", Trends in Food Science & Technology, 2023<br>Publication | <1% |
| 6 | Submitted to Nanyang Technological University<br>Student Paper | <1% |

| | | |
|---|---|---|
| 7 | Submitted to University of Nizwa<br>Student Paper | <1% |
| 8 | discovery.researcher.life<br>Internet Source | <1% |
| 9 | download.bibis.ir<br>Internet Source | <1% |

| | | | |
|---|---|---|---|
| Exclude quotes | Off | Exclude matches | Off |
| Exclude bibliography | On | | |