ISTP

RESEARCH ARTICLE

AI Translation: Evaluating ChatGPT's Reliability in Translating Arabic to English

Mohammed Nour Abu Guba[1] **and Abdallah Abu Qub'a**[2]

[1]English Language Institute, University of Sharjah, Sharjah, the UAE

[2]College of Arts, English Language Department, King Faisal University, Alhasa, Saudi Arabia

***Abstract*:** Machine translation has undergone remarkable evolution since its early rule-based systems in the 1950s, progressing through statistical models in the 1990s to neural machine translation (NMT) in the 2010s. The introduction of large language models, such as OpenAI's GPT series, has marked a new era in translation technology, enabling systems to understand context, tone, and meaning beyond literal word substitution. These developments have reshaped translation research and practice, especially in academic and professional settings. This article explores the effectiveness of ChatGPT as a translation tool in academic contexts, particularly within the fields of humanities and social sciences. Drawing on recent literature, the study reviews advances in prompt engineering, comparative evaluations with traditional machine translation systems, and domain-specific translation challenges. Structured prompts are shown to significantly enhance translation accuracy, with BLEU scores improving as prompt complexity increases. Comparative studies reveal that ChatGPT generally produces more fluent and contextually appropriate translations than tools like Google Translate, especially for high-resource languages and conversational texts. However, its performance declines with specialized terminology, low-resource languages, and culturally embedded expressions. Results show that ChatGPT can be a reliable translation tool that captures the intended meaning rather than offering word-for-word translations, making it a valuable resource in academic and professional settings. Nonetheless, challenges remain, particularly in accurately translating culture-bound expressions, technical jargon, and dialectical variations. Examples from Arabic–English translations underscore these limitations, highlighting instances where ChatGPT succeeded in conveying nuanced meaning and others where it produces awkward or inaccurate renderings. The study concludes by emphasizing the need for ongoing refinement in prompt design and hybrid human–AI translation approaches to enhance translation quality and cultural sensitivity in academic discourse.

***Keywords*:** academic translation; ChatGPT; culture-bound expressions; prompt engineering; translation quality

I. INTRODUCTION

Translation is more than a linguistic process; it is an intricate act of cultural mediation that requires a deep understanding of both source and target languages. With rapid integration and advancements in artificial intelligence (AI) [1], machine translation (MT) tools, such as ChatGPT, have emerged as powerful alternatives to traditional translation methods. While AI-based translators offer speed and accessibility, their effectiveness in preserving linguistic accuracy and cultural aspects remains a subject of debate [2]. Arabic–English translation, in particular, presents unique challenges due to the structural and semantic complexities of Arabic and the cultural distinctions embedded in both languages [3].

The ability of ChatGPT to translate Arabic text into English is crucial for researchers, educators, and professionals in all fields. Prior studies on MT have highlighted improvements in fluency and grammatical accuracy but also revealed persistent issues in handling idiomatic expressions, context-dependent meanings, and culturally specific terms [4]. One of the core challenges in translating Arabic into English lies in preserving the depth of meaning, particularly in texts with religious, rhetorical, or philosophical content.

AI has become an integral part of every aspect of our life including academia [5,6]. Many people rely on AI tools to translate texts [7]. Therefore, it is essential to assess its performance in real-world applications, particularly in academic and professional settings where precision is paramount. This will provide insights into the extent to which AI translation tools like ChatGPT succeed in capturing the intricacies of Arabic discourse.

More specifically, this study aims to evaluate ChatGPT's capabilities in translating Arabic texts to English by analyzing its linguistic accuracy, contextual appropriateness, and ability to convey cultural meanings. Through a systematic assessment of translations, the research seeks to determine ChatGPT's strengths and limitations, offering insights into its potential as a tool for translators, language learners, and cross-cultural communication. The findings will contribute to ongoing discussions about the role of AI in translation and provide recommendations for improving reliability in linguistic and cultural contexts.

The rest of the paper is structured as follows. Section II reviews previous work on using AI as a translation tool and establishes the gap in the literature. Section III lays out the methodology used to collect the data. Section IV presents and discusses the results of ChatGPT's translation, and Section VI discusses professional

---

Corresponding author: Abdallah Abu Quba (e-mail: aabuquba@kfu.edu.sa)

1

translators' opinion of ChatGPT's translation. We conclude the study with some recommendations in Section VII.

II. LITERATURE REVIEW

The translation of AI tools such as ChatGPT has triggered significant interest regarding their capabilities and limitations across languages. In this review, we summarize the main findings and establish the gap in the literature.

The performance of AI tools in translating general texts has been found to be somewhat satisfactory. Early evaluations [8] established that ChatGPT, particularly when prompted carefully, performs competitively with commercial systems such as Google Translate in high-resource European languages. However, its performance substantially lagged for low-resource or distant language pairs. Similarly [9], reported that ChatGPT is highly appreciated for its efficiency and accuracy in translation tasks and emphasized that simplistic prompting often underutilized ChatGPT's capabilities; they proposed enhanced prompt designs —specifically Task-Specific and Domain-Specific Prompts— which significantly improved translation quality. Likewise [10], assessed ChatGPT's translation of 50 Turkish texts into English in the field of education. They confirmed that ChatGPT could be an effective, reliable, dependent translator. Reference [11] also demonstrated that incorporating translation task information, domain specificity, and Part-of-Speech tagging into prompts could substantially enhance translation outputs, confirming that prompt engineering was a critical factor in optimizing ChatGPT's translation effectiveness. In a focused comparison study, ChatGPT was benchmarked against human translators and Google Translate in translating English to Mandarin [12]. Results showed no statistically significant quality difference among the three methods, suggesting ChatGPT's translations could be virtually indistinguishable from human output under certain conditions. However, the small scale of the study was noted as a limitation, pointing to the need for broader validation. Building on these findings [8], highlighted the significant improvement achieved with the transition to GPT-4, noting a notable reduction in hallucinations and mistranslation errors. Their introduction of "pivot prompting"— translating first into a high-resource language—further boosted performance in distant language translations, positioning ChatGPT as a genuinely competitive MT system across a wider array of languages.

In the same vein, AI tools had limitations with low-resource languages, unlike widely spoken languages, most probably due to the scarcity of training data. Reference [13] evaluated ChatGPT-4 alongside Claude-3 and Palm-2 for Saudi Arabic translation and concluded that ChatGPT outperformed all other models but still struggled with the nuanced semantic richness of dialectal Arabic.

Studies comparing ChatGPT with other translation tools such as Google Translate yielded mixed results. Reference [8] found that while ChatGPT performed competitively with commercial translation systems for high-resource European languages, its performance declined for low-resource and linguistically distant languages. Similarly, studies targeting Arabic and English translations were not very common. Reference [14] assessed the accuracy of AI tools in translating Arabic research titles into English, comparing Google Translate, Gemini, and ChatGPT. Their findings revealed that while Gemini produced the least errors, human translations still outperformed AI-generated translations in terms of equivalence and diction accuracy. The study also highlighted the prevalence of polysemous term mistranslations and syntax errors in

AI translations, underscoring the need for improvements in AI-based translation technology. References [15,16] showed that AI-based translation tools often struggled with domain-specific terminology. Reference [15] examined ChatGPT's performance across various genres, including legal, medical, and literary texts, and found that while it handled general texts well, it struggled with technical terminology and cultural nuances. Reference [17] conducted an error analysis of scientific text translations from English to Arabic, revealing that Google Translate outperformed ChatGPT in accuracy, particularly in handling specialized terminology and maintaining textual coherence.

Reference [18] observed that ChatGPT's Arabic–English translations were generally more natural than those of Google Translate but still required minor adjustments. Reference [19] confirmed these findings, noting that while ChatGPT slightly outperformed Google Translate in translating short, contextualized sentences, it struggled with domain-specific texts, emphasizing the need for human oversight. Reference [20] extended this analysis to Arabic dialects, revealing that ChatGPT and Bard outperformed Google Translate for certain dialects but remained inconsistent in handling Classical Arabic and Modern Standard Arabic.

Focusing on translating metaphorical texts, [21] analyzed how Google Translate, ChatGPT, and Gemini translated Arabic idioms into English. They found that literal translation was very common across the three tools, concluding that these tools still needed significant improvement in translating non-literal language.

Results of the performance of AI in translating specialized content showed that these tools still had problems in the domain of literary translation. To recap, while ChatGPT demonstrated promising capabilities in MT, it had limitations in handling complex tasks and could exhibit biases and inaccuracies. More research on the translation of Arabic texts in specific fields is still needed. In this paper, we attempt to assess ChatGPT's performance in translating Arabic texts into English in the fields of humanity and social sciences.

More specifically, the study aims to:

- Assess the accuracy of ChatGPT's Arabic-to-English translations in humanity and social sciences fields. This will shed light on the extent to which ChatGPT can be a reliable tool in translating Arabic texts into English.

- Identify ChatGPT's strengths and weaknesses in translating Arabic texts into English in humanity and social sciences fields.

- Evaluate ChatGPT-generated translations as perceived by professional translators.

By analyzing the strengths and weaknesses of ChatGPT in translating domain-specific texts (which can be a good testing ground), this study contributes to a deeper understanding of the translation potential of ChatGPT in particular and AI tools in general, which would ultimately help improve the quality of AI translation.

III. METHODOLOGY

This study employed a qualitative analysis to evaluate ChatGPT's effectiveness in translating Arabic texts into English across the fields of social sciences and humanities. These categories were chosen to assess the model's ability to handle different linguistic complexities, cultural nuances, and stylistic features. By examining ChatGPT's performance in diverse textual contexts, this research aims to provide a well-rounded evaluation of its translation strengths and limitations.

The study analyzed a total of 15 randomly selected Arabic texts that belong to the fields of sociology (five texts), linguistics (five texts), and education (five texts).

Each of the 15 selected texts was input into ChatGPT using a standardized prompt requesting a non-literal English translation. No additional context, clarification, or human intervention was provided to ensure that the output reflected the AI model's raw translation capabilities. This approach allowed for an objective analysis of how ChatGPT processes Arabic-to-English translation without post-editing, ensuring a fair and consistent evaluation of its strengths and weaknesses.

3.1 EVALUATION CRITERIA

The translations were assessed based on four key factors: semantic equivalence, linguistic accuracy, cultural appropriateness, and technical precision.

- **Semantic equivalence** measured how faithfully and accurately the translation preserved the original meaning and the intended message.
- **Linguistic accuracy** evaluated the readability, grammatical correctness, and naturalness of the English output, ensuring that the translated text conformed to standard English conventions.
- **Cultural appropriateness** examined the retention of rhetorical devices, idiomatic expressions, and culturally embedded references, assessing how well ChatGPT maintained the intricacies of Arabic texts in translation.
- **Technical precision** focused on the handling of specialized terminology and discipline-specific jargon.

By applying these criteria, the study aimed to provide a detailed evaluation of ChatGPT's translation performance across different textual genres, highlighting both its capabilities and areas for improvement.

To further check ChatGPT's translation, five professional translators were requested to evaluate the translations produced by ChatGPT. They were not informed that the texts were AI-generated, allowing for an impartial review based solely on linguistic quality. Each translator assessed the translations using a scoring system from 1 to 10, focusing on four key aspects: semantic equivalence, linguistic accuracy, cultural appropriateness, and technical precision (see Section V for details).

IV. RESULTS AND DISCUSSION

The evaluation of ChatGPT's translation performance across social sciences and humanities reveals notable strengths and weaknesses. ChatGPT demonstrates fluency and coherence in the transfer of the general meaning but faces challenges in handling specialized terminology, syntactic complexity, and cultural subtleties.

ChatGPT demonstrates strong performance in translating straightforward declarative statements. Its translation is clear and readable with a logical flow and sentence structure. Consider the following example:

Original Arabic: لا نجد لغة تمنح الناطق بها فسحة في التعبير عن المعاني بأساليب شتى

ChatGPT's Translation: We find no language that grants its speakers such a wide scope for expressing meanings in various ways.

This translation shows that ChatGPT is able to produce an accurate translation semantically and grammatically. It can convey the content competently with coherent and smooth language.

Another strength of ChatGPT is its ability to fix typos and mistakes in the source language, which means it is able to understand the context. The examples below show how ChatGPT is able to deal with the errors in the source text and translate them correctly.

(1) Mistakes in the source text that are translated correctly (the mistakes are underlined):

a. وإفد أظهرت نتائج التحليل توافر درجة عالية من الأنشطة الواقعية لمهارة القراءة <في المنهاج> The analysis results indicated a high availability of real-life reading activities in the curriculum.

b. <وتمكين الطبة باستخدام اللغة الأنجليزية في مجال الأكاديمي and enhanced their academic use of the language.

c. اراء ايجابية في تبني هذه الطريقة في ظل توفر الظروف لذلك. أما الفريق الثاني؛ رفضوا الفكرة. ... > Some instructors viewed virtual assessment positively, provided that the necessary conditions were met. Others, however, opposed the approach due to concerns over question design.

d. وإذا رمنا ذلك بصدق فلزام علينا أن نصل أنفسنا بيئة الخطاب الأول الذي اعتاده العربي الأول If we sincerely seek this understanding, we must immerse ourselves in the linguistic environment of the first Arabs, who were accustomed to the original discourse.

In (a), the particle "إفد" was misspelled, but ChatGPT is able to understand the context and translate it correctly. In (b), the definite article is missing, but ChatGPT renders it correctly in the target language. In (c), the use of the semicolon is wrong in the source language. However, ChatGPT does not struggle with the resulting fragment and is able to translate it correctly. Finally, in (d), a preposition is missing in the source language, rendering it ungrammatical in Arabic; nevertheless, ChatGPT is able to understand the context and yield a meaningful translation. These observations confirm that ChatGPT is able to handle language problems in the source text by understanding the context.

On the other hand, ChatGPT encounters challenges when handling complex syntactic structures, idiomatic, and culture-bound expressions.

Below are some illustrative examples of problematic translations.

(1) **Original Arabic:** النبي العربي الأمي
- **ChatGPT's Translation:** The Arabic unlettered Prophet
- **Corrected Translation:** The Arab unlettered Prophet

The error here stems from the fact that "Arabic" refers to the language, while "Arab" is the correct term for a person. This distinction is particularly important in religious and historical contexts.

Another example is the mistranslation of transitional phrases and discourse markers, as shown in (2).

(2) **Original Arabic:** وبعد
- **ChatGPT's Translation:** *To proceed*
- **Corrected Translation:** *moving on, so*, or *now then*

This phrase is a classical Arabic rhetorical device used to introduce the main subject following an introduction, and "To proceed" does not fully capture its intended meaning in English. Such expressions are culture-bound and cause difficulties in translation [22,23].

Another problem related to inconsistency in term use. ChatGPT sometimes does not use specialized terms consistently. For example, ChatGPT correctly used the word "acquisition" to translate the Arabic term "اكتساب". However, when the source text used a synonym of that word in a later mention in the text, ChatGPT used a different word (namely performance). If this was translated

by a human translator, it would keep the same term, as is the norm in academic fields. This mistake is triggered by the source text, as will be demonstrated below.

Sometimes, the source text is translated correctly; nevertheless, it lacks clarity. Consider the example below.

(3) **Original Arabic** اذ يأتي اكتساب المقاطع النحوية في اللغة الإنجليزية بتسلسل متوقع. وبالرغم من ذلك، بالنسبة للمتعلمين العراقيين، فإن الحصول على المقاطع النحوية باللغة الإنجليزية قد يمثل تحدياً بسبب الاختلافات في بنية لغتهم الأم

ChatGPT's Translation: The acquisition of grammatical morphemes in English follows a predictable sequence. However, for Iraqi learners, acquiring English grammatical morphemes can be challenging due to differences in the structure of their native language.

It is understood that the writer means "due to the differences between the learners' mother tongue and the target language" (This is clear because the writer talks about the differences between L1 and L2 later in the text). However, ChatGPT is unable to understand the context enough to be able to know that this is the intended meaning.

A few grammatical mistakes are found. These represent advanced grammatical issues such as dangling modifiers, as illustrated in (4).

(4) **Original Arabic:** من أجل ذلك؛ قمنا بإرسال استبيان الكتروني و كشفت الاجابات عن وجود تضارب في الاراء. اراء ايجابية في تبني هذه الطريقة في ظل توفر الظروف

ChatGPT's Translation: To achieve this, an online questionnaire was distributed, revealing conflicting opinions. Some instructors viewed virtual assessment positively, provided that the necessary conditions were met.

This translation has a grammatical mistake. It results in a dangling modifier. There should be a subject for the verb "To achieve." So, it should be: To achieve this, I (or we or the researcher/s) distributed . . . .

Some translations lack precision in technical terminology, which occasionally leads to misrepresentation of academic concepts. For example, ChatGPT translates "الانتماء العرقي" as "racial affiliation," which, while technically correct, is less commonly used in academic discourse than "ethnic identity." Similarly, "الهيمنة الثقافية" is translated as "cultural control" rather than the more precise "cultural hegemony," thereby losing some of the critical meaning associated with Gramscian theory. Another example relates to oversimplifying theoretical concepts, as in "النمط المعياري", which is translated as "standard pattern," which does not fully capture the nuance of "normative framework" in sociological contexts. This tendency to simplify intricate ideas could lead to misunderstandings in academic discussions. This is in line with [17] who emphasized that AI models occasionally misinterpret Arabic scientific and academic terminology, leading to errors in structural coherence.

ChatGPT is sometimes unable to detect and translate figurative language appropriately, which is an obstacle to human translators, too [24,26]. It sometimes translates metaphors, idioms, and rhetorical devices literally, stripping the text of its intended expressive force, which lends support to earlier research, for example, [27,28] who confirmed that ChatGPT struggles with Arabic proverb and idiom translation, often defaulting to literal translations that fail to convey intended meanings.

The following example in (5) illustrates this.

(5) Literal translation of metaphors "نهر الزمن" > "the river of time,"

This translation is a direct translation that does not reflect the conventional English phrase "the passage of time."

Another issue in ChatGPT's translation is the literal translation of technical terms. Phrases such as "تدخل على الفعل المضارع"

(translated as) "enters the present verb," which is grammatically correct but unnatural in English. A better rendering would be "prefixed to a present-tense verb." This finding is in line with [15,16] findings that AI-based translation tools often struggle with domain-specific terminology.

Some inaccurate renderings are source-text-induced because the source text is not well written or defined. Consider the translations in (6).

(6) Source-text induced issues

a. فإن العلم بالعربية لغة القرآن فريضة من فرائض تبنى عليه الأحكام الشرعية ، (والفهم الصحيح لمراد الله تعالى (بألفاظه. وتراكيبه
The knowledge of Arabic, the language of the Quran, is an essential obligation upon which Islamic rulings and the correct understanding of Allah's intended meanings in <u>His words and structures</u> are based.

b. جمعت البيانات من المتعلمين على مرحلتين: اختبارقبل وبعد الدرس في نهاية الحصص الدراسية من دروس القواعد.
Data was collected from learners in two phases: a pre-test and a post-test conducted at the end of grammar lessons.

c. اثرت جائحة الكورونا على التعلم و اختبار المتعلمين.

The COVID-19 pandemic significantly affected learning and student assessment.

In (6a), the underlined phrase (بألفاظه. وتراكيبه) is ambiguous as it can modify the word "Alla" syntactically. However, we know, based on our background knowledge, that this modifies the word "Quran."

In (6b), the source text does not use the modifiers correctly and this affected ChatGPT's translation clarity. If this was translated by a human being, it would be rendered with better phrasing as "a pre-test administered before the grammar lessons and a post-test conducted at the end of these lessons."

In (6c), the word "student" modifies both "learning" and "assessment" and therefore should be translated as "students' learning and assessment."

Such cases show that the Arabic text to be translated should be written clearly with no ambiguity that might cause different interpretations. This means one should edit their text before submitting it to ChatGPT; otherwise, such problems may arise.

V. EXPERT TRANSLATORS' REVIEW AND RECOMMENDATIONS

The results of the evaluation of the five professional translators show that ChatGPT's translation is generally accurate and accepted. Table I presents the translators' scores of ChatGPT's translation.

In terms of semantic accuracy, it received an average of 9.4 out of 10, which means there are very minor problems with translating content. With respect to grammatical accuracy, translators concurred that it is excellent, with an average at 9.8/10. Regarding cultural appropriateness and technical precision, ChatGPT received a lower score at 8.4/10 for both. This is in conformity with our evaluation that confirms that the only real problem with ChatGPT's translation lies in its dealing with culture-bound and technical terms. The translators highlighted almost the same issues that were dealt with us above, for example, racial affiliation instead of "ethnic identity." They reiterated that some terms were not the best fit in academic discourse, and they suggested using terminology that reflects standard usage in academic publications, as in "enters the present verb" for "is prefixed to a present-tense verb." Some comments were related to style, which cannot be considered

Table I. Translators' scores of ChatGPT's translation

	Translator 1	Translator 2	Translator 3	Translator 4	Translator 5	Average
Semantic accuracy	9	10	9	9	10	9.4
Grammatical accuracy	9	10	10	10	10	9.8
Cultural appropriateness	8	9	8	8	9	8.4
Technical precision	8	9	8	9	8	8.4
Overall average	8.5	9.5	8.75	9	9.25	9

mistakes but rather reflect personal preferences of language use. For example, it was suggested that "the study's findings" be replaced by "the findings of the study" and "various" by "a variety of."

Taken together, the results of our analysis and the views of the professional translators show that ChatGPT can be an excellent translation tool that is able to produce grammatically coherent and readable sentences that are fully accessible to general readers, which is in line with previous studies [29,30]. However, it sometimes struggles with technical terms and metaphorical expressions, which is in line with previous studies that confirmed that such expressions caused problems for AI tools [15].

VI. CONCLUSION AND RECOMMENDATIONS

The overall performance of ChatGPT in translating social sciences and humanities texts demonstrated its strengths in linguistic and content accuracy. It used academic and correct English fluently with accurate sentence structure and a high level of readability. It was able to domesticate the translation to produce natural idiomatic English. It was often able to understand the context and fix typos and mistakes in the source text. These aspects constitute ChatGPT's capabilities. These results indicated that ChatGPT could be a useful translation tool and can generally be a reliable translator. However, ChatGPT still had some limitations. Its outputs required human refinement and post-editing to ensure quality and precision in its use of technical terms and cultural appropriateness. This means that there is still room for improving AI tools (cf. [14,31,32], where more focus should be on greater contextual awareness, flexibility in handling complex arguments, and enhanced recognition of figurative and specialized language.

These results corroborated previous research that reported that AI translation often struggled with idiomatic expressions and contextual accuracy.

In line with these results, the following recommendations are in order. ChatGPT should be provided with a refinement system that tags technical terms to allow for a specialized translation within a certain field, which could significantly enhance accuracy. The same applies to idiomatic expressions. If metaphors are tagged in source corpora and AI is trained to detect these phrases in order not to translate them literally, a more accurate translation will be produced. Moreover, a reader-adaptive translation mode could be a valuable feature for ChatGPT. This would allow users to select whether they want a translation geared toward specialists, general academics, or the public. The AI could then adjust terminology, sentence structure, and explanatory depth accordingly. In the same vein, prompting should be specific enough to yield more accurate and contextually relevant translations. This may include translation task information and context domain information. Finally, it is highly recommended that the Arabic text to be translated should be written clearly with no ambiguity that might cause different interpretations.

Future research should focus on refining prompt engineering techniques, expanding training datasets, and integrating AI-assisted translation into hybrid models that combine human expertise with machine learning advancements.

AUTHORS' CONTRIBUTIONS

All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

CONFLICT OF INTEREST STATEMENT

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

[1] A. Qub'a et al., "Students' reactions to using ChatGPT in English writing," *International Journal of Innovative Research and Scientific Studies*, vol. 8, no. 1, pp. 2732–2740, 2025.

[2] A. Toral and A. Way, "What level of quality can neural machine translation attain on literary text?" Translation Quality Assessment, from Principles to Practice, Springer, ArXiv, abs/1801.04962, 2018, pp. 263–287. https://link.springer.com/book/10.1007/978-3-319-91241-7

[3] N. Habash, Introduction to Arabic Natural Language Processing. Morgan & Claypool Publishers, 2010, pp. 157–158. https://link.springer.com/book/10.1007/978-3-031-02139-8

[4] R. Seyidov, "Importance and role of artificial intelligence for Arabic translators," *International Journal for Arabic Linguistics & Literature Studies (JALLS)*, vol. 6, no. 2, pp. 71–83. 2024. https://doi.org/10.31559/JALLS2024.6.2.4

[5] B. Lainjo, From Data Analysis to Creative Arts: The Ubiquity and Impact of Artificial Intelligence in Academia. *Int. J. Educ. Teach. Soc. Sci.*, vol. 4, no. 4, pp. 55–73, 2024.

[6] H. Crompton and D. Song, "The potential of artificial intelligence in higher education," *Rev. Virtual Univ. Catol. Norte*, vol. 62, pp. 1–4, 2021.

[7] M. Farghal and A. Haider, "Translating classical Arabic verse: human translation vs. AI large language models (Gemini and ChatGPT)," *Cogent Soc. Sci.*, vol. 10, no. 1, pp. 1–15, 2024.

[8] W. Jiao et al., Is ChatGPT a good translator? A preliminary study. arXiv preprint, arXiv:2302.08092. 2023. https://doi.org/10.48550/arXiv.2302.08092

[9] K. Peng et al., Towards Making the Most of ChatGPT for Machine Translation: Findings of the Association for Computational Linguistics: EMNLP 2023, 2023. 2023. https://www.scopus.com/record/display.uri?eid=2-s2.0-85182577149&origin=scopusAI

[10] Ç. Işım and Y. Balcioglu, ChatGPT: Performance of translate [Paper presentation]. Paper presented at the 3rd International Acharaka Congress on humanities and Social Sciences, Acharaca. 2023.

[11] Y. Gao, R. Wang, and F. Hou, Unleashing the Power of ChatGPT for Translation: An Empirical Study. 2. 2023. https://doi.org/10.48550/arXiv.2304.02182

[12] C. Woodrum, "ChatGPT and Language Translation," in *Artificial Intelligence in HCI*, H. Degen, S. Ntoa Eds., HCII 2024. Lecture Notes in Computer Science, vol. 14736. Cham: Springer, 2024, pp. 143–156. https://doi.org/10.1007/978-3-031-60615-1_10

[13] E. A. Alghamdi, J. Zakraoui, and F. A. Abanmy, Domain Adaptation for Arabic Machine Translation: The Case of Financial Texts. *ArXiv, abs/2309.12863.* 2023.

[14] S. Al-Salman and A. S. Haider, "Assessing the accuracy of MT and AI tools in translating humanities or social sciences Arabic research titles into English: Evidence from Google Translate, Gemini, and ChatGPT," *International Journal of Data and Network Science*, vol. 8, no. 4, pp. 2483–2498, 2024.

[15] F. Khoshafah, "ChatGPT for Arabic-english translation: Evaluating the accuracy," *Res. Sq.*, 2023. https://doi.org/10.21203/rs.3.rs-2814154/v2.

[16] R. Al-Jarf, "Translation of medical terms by AI: A comparative linguistic study of microsoft copilot and google translate," in *Artificial Intelligence and Its Practical Applications in the Digital Economy*, Y. M. Elhadj, M. F. Nanne, A. Koubaa, F. Meziane, and M. Deriche Ed., I2COMSAPP 2024. Lecture Notes in Networks and Systems, vol 862. Cham: Springer, 2024, pp. 239–251. https://doi.org/10.1007/978-3-031-71429-0_17

[17] E. Alzain, K. A. Nagi, and F. Algobaei, "The quality of Google Translate and ChatGPT english to Arabic translation: The case of scientific text translation," *Forum Linguist. Stud.*, vol. 6, no. 4, pp. 837–849, 2024.

[18] M. A. AlAfnan, "Large language models as computational linguistics tools: A comparative analysis of ChatGPT and Google machine translations," *J. Artif. Intell. Technol.*, vol. 5, pp. 20–32, 2024.

[19] L. Alkhawaja, "Unveiling the new frontier: ChatGPT-3 powered translation for Arabic-English language pairs," *Theory Pract. Lang. Stud.*, vol. 14, no. 2, pp. 347–357, 2024.

[20] K. Kadaoui et al., TARJAMAT: Evaluation of Bard and ChatGPT on machine translation of ten Arabic varieties. arXiv preprint, 2023. arXiv:2308.03051. https://doi.org/10.48550/arXiv.2308.03051

[21] M. M. Obeidat et al., "Analysing the performance of gemini, ChatGPT, and google translate in rendering english idioms into Arabic," *FWU J. Soc. Sci.*, vol. 18, no. 4, pp. 1–18, 2024.

[22] A. Al-Hassan, "The importance of culture in translation: Should culture be translated?" *Int. J. Appl. Linguist. Engl. Lit*, vol. 2, no. 2, pp. 219–225, 2013. https://www.scopus.com/record/display.uri?eid=2-s2.0-84946400311&origin=scopusAI

[23] A. M. S. Al-Hamzi et al., "Socio-pragmatic analysis of utterances with polite addressing terms: Translation shift across Arabic-english cultures," *Cogent Arts Humanit.*, vol. 11, no. 1, pp. 2341994, 2024. https://www.scopus.com/record/display.uri?eid=2-s2.0-85197736202&origin=scopusAI

[24] A.-F. Abu-Ssaydeh, "Translation of english idioms into Arabic," *Babel. Revue internationale de la traduction /International Journal of Translation*, vol. 50, no. 2, pp. 114–131, 2004.

[25] L. Chen, "On literal translation of english idioms," *Engl. Lang. Teach.*, vol. 2, no. 2, pp. 164–166, 2009.

[26] K. A. ElFqih and J. Monti, "Large language models as legal translators of Arabic legislation: Do ChatGPT and Gemini care for context and terminology?" In N. Habash, S. Khalifa, M. Abdul-Mageed, A. Abdelali, , H. Sajjad, & F. Bougares Eds., *Proceedings of the Second Arabic Natural Language Processing Conference* (ArabicNLP 2024), Association for Computational Linguistics, 2024, pp. 111–122.

[27] B. Tannous,Assessing the Accuracy of AI Tools (Google Translate, ChatGPT, and Gemini) in Translating Arabic Idiomatic Expressions and Proverbs into English [MA thesis]. Jordan: Applied Science Private University, 2024.

[28] M. El-Saadany, "A comparative study between Chat GPT and human translation in translating english proverbs into Arabic," *J. Sci. Res. Arts*, vol. 25, no. 5, pp. 24–54, 2024.

[29] R. Alkhatib and A. S. Haider, "Subtitling legal expressions in English series into Arabic by Netflix, machine, and artificial intelligence," *Pak. J. Criminol*, vol. 14, no. 4, pp. 513–528, 2024.

[30] J. Wang, "Exploring the potential of ChatGPt-4o in translation quality assessment," *J. Theory Pract. Humanit. Soc. Sci.*, vol. 1, no. 3, pp. 18–30, 2024. https://woodyinternational.com/index.php/jtphss/article/view/27

[31] S. Öner Bulut and N. Alimen, "Translator education as a collaborative quest for insights into the re-positioning of the human translator (educator) in the age of machine translation: The results of a learning experiment," *Interpret. Transl. Train.*, vol. 17, no. 3, pp. 375–392, 2023.

[32] A. A. Qub'a, M. N. A. Guba, and S. Fareh, "Exploring the use of grammarly in assessing english academic writing," *Heliyon*, vol. 10, no. 15, p. e34893, 2024.