Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Special Issue on ChatGPT, April 2024 Pp. 237-251
DOI: https://dx.doi.org/10.24093/awej/ChatGPT.16

Evaluating Translation Tools: Google Translate, Bing Translator, and Bing AI on Arabic Colloquialisms

Hamad Abdullah H Aldawsari
Department of English, College of Sciences and Humanities
Prince Sattam Bin AbdulAziz University, Wadi ad-Dawasir, Saudi Arabia
Email: h.alnawwar@psau.edu.sa

Received:01/25/2024 Accepted:03/26/2024 Published: 04/24/2024

Abstract
This study examines the advancements in AI-driven machine translation, specifically focusing on the accurate translation of Arabic colloquial expressions. It aims to assess the progress made by Large Language Models, such as Bing AI Chat, compared to traditional machine translation systems. By focusing on colloquial expressions, this research aims to shed light on the challenges and opportunities for improvement in machine translation systems, particularly when dealing with the complexities of translating informal Arabic utterances. Building upon At-tall's 2019 thesis, which compared Google Translate and human translators, the study employs the same Arabic sentences as a test dataset, allowing for a direct comparison between 2019 translations and those produced by current machine translation tools. The findings indicate limited improvement in Google Translate since 2019, with Bing Translator exhibiting a similar level of translation accuracy. In contrast, Bing AI Chat consistently outperformed the other systems, showcasing the potential of Large Language Model machine translation. Notably, Bing AI Chat provided interpretations and valuable comments on the tested Arabic phrases, demonstrating a deeper understanding of the intended meaning. This study contributes significantly to the field of machine translation by providing evidence of the potential of Large Language Model systems in producing more accurate Arabic-English translations. It emphasizes the advantage of Large Language Models in dealing with non-standard Arabic expressions, encouraging further exploration of Large Language Model-powered approaches in machine translation. The findings offer a promising pathway towards achieving more accurate and expressive translations across diverse languages and cultures.

Keywords: Arabic colloquial expressions, applied linguistics, Large Language Model systems, machine translation, Bing Translator, Bing AI Chat

Cite as: Aldawsari, H. A. (2024). Evaluating Translation Tools: Google Translate, Bing Translator, and Bing AI on Arabic Colloquialisms. *Arab World English Journal (AWEJ)* Special Issue on ChatGPT, April 2024: 237-251.DOI: https://dx.doi.org/10.24093/awej/ChatGPT.16

Introduction

In recent years, machine translation systems have undergone significant advancements, fueled by the rapid progress in Artificial Intelligence (AI) technologies (See Almahasees, 2021). This research paper aims to compare the translation output of three prominent machine translation systems: Google Translate (GT), Bing Translator (BT), and Bing AI Chat (BAI). Specifically, it seeks to investigate whether translation quality has improved from 2019 to the present day. Additionally, the study will examine whether BAI, an AI-powered Chatbot, provides more accurate translations compared to traditional machine translation websites like GT and BT.

To establish a baseline for this analysis, this study will revisit a 2019 thesis titled "Comparative Study Between Google Translator and Human Translator in Rendering Colloquial Arabic Expressions in the Late Prime Minister Wasfi At-Tall's Speeches into English" by Shaima'a Mohammad At-tall (At-tall, 2019). The author's findings revealed poor performance by machine translations in rendering colloquial Arabic expressions accurately, highlighting the continued challenges in this complex domain (At-tall, 2019).

The first objective of this research is to assess whether there have been notable improvements in translation output since 2019. By evaluating translations from the aforementioned thesis and comparing them with the current outputs of GT, BT, and BAI, this study aims to identify any advancements in the accuracy of these Machine Translation (MT) and Large Language Model systems (LLM). The second objective is to examine whether BAI, which utilizes LLM technologies, provides more accurate translations compared to traditional MT systems. The study will compare the translation quality of BAI with that of GT and BT, focusing on factors such as accuracy and readability.

By addressing these objectives, the study aims to provide insights into the progress made in machine translation systems and the potential advantages offered by AI-driven translation technologies. Understanding the strengths and limitations of these systems will contribute to a comprehensive understanding of the current state of machine translation and its implications for cross-lingual communication.

This study holds significant importance as it aims to qualitatively assess the translation capabilities of two MT systems while also assessing the performance of LLM systems in translating Arabic colloquial expressions. By focusing on colloquial expressions, this research aims to shed light on the challenges and opportunities for improvement in these systems when dealing with the complexities of informal Arabic language usage. The findings of this study will contribute valuable insights to the field of MT and LLM informing prospective advancements in accurately translating Arabic colloquial expressions.

As mentioned, the research objectives encompass evaluating the accuracy and quality of translation output by GT, BT, and BAI for Arabic colloquial expressions, as well as comparing their performance across different types of informal language usage. The research questions guiding this study are as follows:

1. How accurately do Google Translate, Bing Translator, and Bing AI Chat translate Arabic colloquial expressions?
2. What are the advancements and variations in performance among these systems in translating Arabic colloquial expressions?

Overall, this research will provide valuable insights into the translation of Arabic colloquial expressions and contribute to the ongoing efforts to improve the accuracy and quality of MT and LLM systems. By comparing the translation output of GT, BT, and BAI, this study seeks to

evaluate whether LLM offers more accurate translations compared to traditional machine translation websites. By analyzing the findings, the research aims to shed light on the advancements in machine translation systems and their implications for effective language translation.

Literature Review

Machine Translation (MT) systems have made significant progress over the years, as evidenced by numerous research studies assessing their performance (Al-Jarf, 2023; Hutchins, 2001; Hutchin & Somers, 1992; Okpor, 2014). However, when it comes to translating Arabic everyday expressions, the results remain mixed (Ali, 2020; Almahasees, 2021; Ameur, Mezaine, & Guessoum, 2020; At-tall, 2019; Banimelhem & Amayreh, 2023; Harrat, Meftouh & Smaili, 2019; Zakraoui, Saleh, Al-Maadeed & Alja'am, 2021). Most significantly, At-tall (2019) conducted a thesis comparing the performance of Google Translator and human translators in rendering colloquial Arabic expressions from the late Prime Minister Wasfi At-Tall's speeches into English. The findings indicated that MT systems produced "low quality" output (p. 42) and performed poorly in accurately conveying colloquial Arabic expressions, highlighting the persistent challenges in this complex domain. Although some recent advancements have been made in LLM systems, current research suggests they may not outperform traditional MT for all tasks. For example, Banimelhem & Amayreh (2023) found that ChatGPT, a well-known LLM system, performed below average compared to 14 established MT systems in translating standard Arabic sentences. This highlights the need for further research that specifically evaluates the capabilities of LLMs in handling colloquial Arabic expressions.

This highlights the need for further investigation into the translation of Arabic colloquial expressions and the difficulties faced by MT and LLM systems. By revisiting At-tall's (2019) thesis, this study aims to gain insights into the performance of machine translations in accurately rendering colloquial Arabic expressions. The research will contribute to a comprehensive understanding of the challenges that persist in this domain and the limitations of MT systems.

Machine Translation

MT has been a subject of extensive research and development since the mid-20th century. It involves the use of computer algorithms to automatically translate text or speech from one language to another (Hutchin & Somers, 1992; Okpor, 2014). Traditional MT systems primarily relied on rule-based or statistical approaches, which often struggled with understanding the complexities of the human language (Hutchins, 2001).

Google Translate and Bing Translator are two widely used MT systems that have gained popularity over the years. GT, developed by Google, and BT, developed by Microsoft, have become go-to tools for many individuals seeking quick and accessible translations (Almahasees, 2021). These systems employ statistical machine translation techniques and utilize vast amounts of parallel bilingual and multilingual data for training their models (Koehn, 2010; Almahasees, 2021). Despite their widespread use, both GT and BT have faced criticism for their limitations in accurately comprehending the hidden meanings of language, particularly when it comes to translating Arabic idiomatic expressions and complex linguistic structures (Aldawsari, 2023; Ali, 2020; Almahasees, 2021; Ameur, Mezaine, & Guessoum, 2020; At-tall, 2019; Harrat, Meftouh & Smaili, 2019; Zakraoui, Saleh, Al-Maadeed & Alja'am, 2021). For instance, a study conducted by Aldawsari (2023) focused on the translations of two MT systems, Google Translate and

SYSTRAN, and concluded that both systems "struggled" in accurately translating complex Arabic linguistic features such as homonyms, heteronyms, and polysemes (p. 27). At-tall (2019, p. 41), found that MT tools "can't cope with most of the mentioned terms in the source language" and that they "directly followed the literal meaning". In other words, although both tools were quick, they produced "low quality" output (p. 42). The purpose of this study is to examine the advancements made in MT tools by comparing them to Large Language Model tools (LLM), with a specific focus on Microsoft Bing Chat.

Large Language Models

Large Language Models (LLMs), such as Microsoft Bing AI Chat (BAI), are highly advanced natural language processing models that have been trained on vast amounts of data. They excel in tasks such as text generation, translation, and question answering. They can produce coherent and contextually relevant responses (Huang, Wu, Liang, Wang & Zhao, 2023). LLMs have the potential to revolutionize content creation and translation due to their ability to mimic human language effectively (Lee, 2023). Microsoft Bing Chat is an implementation of LLM technology developed by Microsoft. It enhances chat-based interactions with intelligent suggestions and contextual recommendations (Rinsum, 2023). It also integrates with Bing Search and Office 365, providing convenient assistance and real-time information retrieval.

However, LLMs and BAI have limitations. LLMs may occasionally generate inaccurate or nonsensical responses due to biases in the training data or because of misinformation (Rinsum, 2023). BAI's effectiveness depends on the quality and comprehensiveness of the underlying data sources and search algorithms. It may struggle with complex or ambiguous queries, necessitating continuous improvement and refinement (Huang et al., 2023). Yet with its AI-driven approach, BAI has the potential to provide more accurate and contextually appropriate translations compared to traditional MT systems like the aforementioned Google Translate and Bing Translator. However, the specific performance and capabilities of BAI in translating different languages, including Arabic, need to be explored further. To date, there has been limited exploration of BAI's performance or LLMs in general in translation compared to dedicated MT tools, yielding inconclusive results (Boughorbel, & Hawasly, 2023; Banimelhem & Amayreh, 2023). For example, Banimelhem and Amayreh (2023) found that ChatGPT, a well-known LLM system, performed below average compared to 14 established MT systems. Notably, their study focused on translating standard Arabic sentences, rather than the more complex and culturally-specific field of colloquial Arabic expressions. This highlights the need for further research that specifically evaluates the capabilities of LLMs in handling colloquial Arabic expressions.

Arabic Colloquial Expressions

Arabic is a language with rich variations, including formal and colloquial registers (Al-Jarf, 2023; Habash & Diab, 2012). Colloquial Arabic expressions refer to informal, region-specific linguistic patterns and idiomatic phrases that may not have direct equivalents in other languages (Al-Saidat, 2011; Al-Jarf, 2023). They can vary depending on the region, dialect, and context of the speaker and the listener (al-Btoush, 2014; Al-Saidat, 2011). They can also reflect the culture, history, and humor of the Arabic people (Alnamer & Alnamer, 2018). Translating colloquial Arabic expressions accurately poses significant challenges for MT systems due to their cultural and contextual complexities (Al-Kharabsheh & Yassin, 2017; At-tall, 2019; Habash & Diab, 2012; Jibreel, 2023). For instance, Al-Kharabsheh and Yassin (2017) identified "major problems" in the

English translations of Arabic colloquialisms found in films, such as mistranslations, omissions, and a loss of meaning (p. 26). Similarly, At-tall (2019) compared human translation and Google Translate's ability to render colloquial Arabic expressions used by former Prime Minister Wasfi At-Tall and found that Google Translate struggled with these informal phrases. While these studies highlight limitations in MT technology, it's important to consider the advancements made in MT and LLM systems since 2019. These advancements may have improved MT's ability to handle the complexities of colloquial Arabic, calling for a re-examination of how well these systems perform in this specific domain. To address this gap, the present study aims to investigate the effectiveness of LLM-based machine translation compared to traditional MT systems specifically in translating Arabic colloquial expressions, building upon the research conducted by At-Tall (2019).

Prime Minister Wasfi At-Tall's Speeches

Wasfi At-Tall was a prominent political figure and statesman in Jordan, serving as the Prime Minister on multiple occasions during the 1960s and early 1970s (Susser, 1994; Joyce, 2008; Morris, 2001). Known for his charismatic leadership and powerful oratory skills, At-Tall was revered for his ability to connect with the masses through his speeches. In his public addresses, At-Tall employed a distinctive style of language that encompassed both formal Arabic and colloquial expressions. This unique blend allowed him to effectively communicate and resonate with a diverse audience, transcending social and educational barriers. At-Tall's speeches were characterized by his passionate delivery and persuasive rhetoric, making them memorable and influential (Susser, 1994). His choice of language, including colloquial Arabic expressions, added a touch of authenticity and relatability, enabling him to connect with the people of the time.

At-tall's Study

The thesis by Shaima'a Mohammad At-tall, titled "Comparative Study Between Google Translator and Human Translator in Rendering Colloquial Arabic Expressions in the Late Prime Minister Wasfi At-Tall's Speeches into English," focuses on the translation of colloquial Arabic expressions in the speeches of Prime Minister Wasfi At-Tall. This study sheds light on the difficulties faced by MT systems, specifically Google Translate, in accurately rendering these expressions into English. The findings of At-tall's thesis indicate the limitations of MT, particularly in handling the intricacies of colloquial Arabic expressions. However, it is important to note that the research was conducted in 2019, and since then, advancements in AI-driven translation technologies and LLM, such as BAI, may have addressed some of these challenges.

Following this, the primary objective of this research is to build upon existing knowledge and provide a comprehensive evaluation of the translation output and advancements in MT systems. Specifically, the research will evaluate the translation by MT and LLM, focusing on the performance of prominent systems such as GT and BT. Additionally, the study will explore the emergence of BAI as an LLM-powered translation system and investigate its potential for providing more accurate translations. Furthermore, this research aims to examine the challenges associated with translating Arabic colloquial expressions, which are known for their cultural and contextual complexities. By analyzing the specific study on Prime Minister Wasfi At-Tall's speeches, the study will provide insights into the limitations faced by MT systems, particularly in rendering colloquial Arabic expressions accurately.

Through this analysis, this research endeavors to contribute to a deeper understanding of the progress made in MT, including the advancements in Google Translate, Bing Translator, and

the potential benefits offered by Bing AI Chat. By evaluating the translation output and improvements in these systems, the study aims to provide valuable insights for enhancing cross-lingual communication and addressing the challenges associated with translating colloquial expressions in Arabic and other languages.

Methods

This research focuses on the quality and treatment of colloquial Arabic in MT over time. It adopts a comprehensive approach by combining previous research with new data and multiple MT and LLM tools. Specifically, it builds on At-tall's 2019 thesis, which compared Google Translator and human translators in rendering colloquial Arabic expressions, providing a background for examining MT system performance. By utilizing the Arabic sentences analyzed in At-tall's thesis as the test dataset, the research enables a direct comparison between 2019 translations and those produced currently by the MT tools.

To broaden the research scope and assess different MT approaches, the analysis incorporates not only Google Translate (GT) but also Bing Translator (BT) and LLM-powered Bing AI Chat (BAI). This expansion allows for a comprehensive evaluation of the advantages and disadvantages associated with these different MT systems. To ensure a fair and precise comparison among the three MT tools, the research employs standardized inputs, eliminating potential inconsistencies that may arise from differences in user input or application interface. The output analysis takes a multi-faceted approach, considering factors such as accuracy, fluency and readability. By evaluating these aspects, the research aims to provide a comprehensive understanding of how each MT system handles the intricacies of colloquial Arabic. To gain insights into the evolution of MT capabilities over time, the research compares At-tall's 2019 findings with the current results. This comparative analysis helps identify possible improvements in MT capabilities and contributes valuable insights to the understanding of MT advancements. By employing this comprehensive approach, the research goes beyond a simple comparison of MT tools. It connects with existing work expanding the scope of analysis, and providing valuable insights into the evolving landscape of MT and LLM systems, specifically its ability to handle the complexities of colloquial Arabic.

Arabic Test Suite

Ten Arabic sentences containing colloquial expressions originally analyzed by At-tall (2019) in her thesis "Comparative Study Between Google Translator and Human Translator in Rendering Colloquial Arabic Expressions in the Late Prime Minister Wasfi At-Tall's Speeches into English" were used as the test suite (Appendix A). This selection ensures a focus on At-Tall's unique linguistic style and its representation of Jordanian colloquial Arabic. The chosen expressions encompassed a variety of types, including idioms, proverbs, and informal vocabulary. Examples include: اللي ما بقبل البلد بضيقها البلد ما بتقبله بعزها. (lit. Whoever does not accept the country when it is poor, the country will not accept him when it is rich), ما حدا بموت ناقص عمر (lit. Nobody will die before his time), and اللي خبزه مش من قمحه عاش مذلول (lit. Whoever's bread is not from his wheat, will live humiliated).

Research Instruments

Three MT tools were employed in this study: Google Translate (GT), Bing Translator (BT), and Bing AI Chat (BAI). GT and BT are well-established MT systems known for their user-

friendly interfaces and widespread use in general-purpose translation. Bing AI Chat, on the other hand, represents a novel approach leveraging a LLM approach. The official websites of GT (Translate.Google.com) and BT (Bing.com/translator) were accessed to utilize their standard translation functionalities. Additionally, BAI (Bing.com/chat) was directly queried using the prompt “Translate this please from Arabic to English” to obtain translations without providing any additional context or explanation. GT was chosen as it was the tool used in At-tall’s (2019) study, providing a benchmark for comparison. BT was added to the evaluation because BAI bases its translations upon it, thus allowing for a comprehensive analysis of the performance of both GT, BT and BAI. What makes BAI particularly interesting is its addition of commentary after each translation. These comments serve to explain the intended meaning behind the translated phrases and offer background information regarding the context of use. These helpful comments or explanations provided by BAI were recorded and compiled alongside the standard translations as BAI commentary (Appendix D). This additional data provides a unique perspective on BAI's approach to translating colloquial Arabic expressions and showcases its potential to engage in interpretation and extend the contextual understanding of the translations.

Evaluation

A direct qualitative approach was employed to evaluate the translation quality of GT, BT and BAI compared to translations in At-tall (2019). This allowed for a comprehensive examination of the different translations. The analysis involved systematically examining the translations and carefully comparing them to the corresponding 2019 translations (At-tall, 2019). Any discrepancies or notable observations were thoroughly documented and analyzed to provide insights into the translation quality of each tool. This analysis focused on comparing the performance of BAI with GT and BT, highlighting its strengths and weaknesses in handling Wasfi At-Tall's colloquial Arabic expressions.

This methodology combines a carefully chosen test suite of real-world expressions with a qualitative evaluation process to provide a comparative analysis of the translation quality offered by BAI and established MT systems when faced with the particular challenges of colloquial Arabic.

Findings

Tables Two, three, and Four present the results obtained after running the test suite through both MT systems and BAI (Appendices B-D). The results show that GT largely maintained the same results as in 2019, suggesting limited improvement in the system over time. Similarly, BT's performance closely mirrored that of GT, indicating a comparable level of translation accuracy between the two systems. On the other hand, BAI consistently delivered substantially more accurate translations, indicating the potential progress in LLM machine translation compared to traditional MT systems. Furthermore, BAI provided comments and interpretations of the Arabic phrases, suggesting a deeper understanding of the intended meaning.

When comparing the translations generated by GT with its 2019 output, it was found that the system largely maintained the same results, indicating limited improvement since 2019. However, there were exceptions observed in sentences three and five (Appendix B). In these cases, GT was able to recognize the colloquial phrases "بدكو" (lit. you want) and "بجوز" (lit. maybe), whereas in At-tall (2019, p. 53), GT transliterated them as "Badko" and "By Joe," respectively. On the other hand, BAI consistently delivered substantially more accurate translations. Despite not

ISSN: 2229-9327

being given any specific context or explanations, it provided comments and interpretations of the Arabic phrases, indicating a deeper understanding of the intended meaning and demonstrating the potential progress in machine translation, particularly within LLM approaches.

In sentence eight, for example, both GT and BT omitted the Arabic possessive pronoun 'ه' (lit. his), resulting in a vague translation. However, BAI demonstrated a remarkable alignment with the human professional translation in At-tall (2019) by retaining the possessive pronoun and producing a more accurate translation, thus preserving the intended meaning of the sentence. Additionally, in sentence nine, GT and BT provided different interpretations based on the diacritical marks in the Arabic language. GT used the possessive pronoun "its" to refer to the country's land and people, while BT interpreted it as "his" referring to a 'king.' In contrast, BAI aligned with GT's interpretation, showcasing its ability to understand the intended meaning and produce a translation consistent with the given context. Similarly, in sentence 10, GT inaccurately translated the colloquial word "إلها" (lit. of it) as "god," while BT maintained a literal translation. BAI, on the other hand, demonstrated a more sophisticated approach by rewriting the phrase to make it clearer to the target readers. It contextualized the meaning and presented it in a more readable manner.

What's more, BAI accompanied its translations with explanatory comments. For instance, in the above example, the mention of "Sykes-Picot" in the source text prompted BAI to provide additional information about the Sykes-Picot Agreement as commentary after the translation (Appendix D). This unique capability of providing commentary alongside the translated text by BAI is a significant advantage over traditional MT systems. This feature enhances the translation output by providing contextual information and insights into the underlying meaning and connotations of the translated text. The inclusion of such commentary can be valuable in improving the overall understanding and interpretation of complex or culturally significant expressions, allowing users to gain deeper insights into the translated content.

BAI also showed improvement by referring to previous requests within the same chat session. This feature enhanced the continuity and coherence of the translations (see use of "also" in sentences two and three, Appendix D). Additionally, it added punctuation marks such as full stops and commas, even if they were not present in the source texts (e.g., sentence two). This adjustment seems aimed at improving the readability and flow of the translated sentences. Furthermore, BAI aimed to enhance the translation process through editing and proofreading, making efforts to refine the translations and achieve improved accuracy. An example of this can be observed in sentence seven, where both GT and BT provided the same inaccurate output: "No one dies without age" and "No one dies minus age" for the sentence ما حدا بموت ناقص عمر (lit. Nobody will die minus age), respectively (At-tall, 2019; Appendices B and C). In contrast, BAI demonstrated a better understanding of the intended meaning by translating it as "No one dies before their time" resulting in a more accurate translation of the source text (Appendix D).

These observations highlight a superior performance of BAI compared to GT and BT in terms of accuracy, contextual understanding, and linguistic refinement. BAI's ability to interpret and improve translations, as well as its incorporation of previous chat session information, contribute to its enhanced performance in generating more accurate translations.

Discussion

The findings reveal varying degrees of accuracy among the three systems. BAI consistently produced the most accurate translations of Arabic colloquial expressions. In contrast, GT largely

maintained the same results since 2019, suggesting limited improvement in its ability to handle these informal phrases. BT performance closely mirrored that of GT, indicating a comparable level of accuracy between the two traditional MT systems.

This study aimed to investigate two important aspects of machine translation advancement: improvements since 2019 and the potential of LLM tools like BAI. The results obtained shed light on these aspects and provide valuable insights into the current state of machine translation systems. Overall, this assessment provides valuable insights into the advancements and potential of machine translation systems. The findings suggest that AI-driven approaches, exemplified by BAI, hold promise for improving the accuracy and quality of translations, especially when handling complex linguistic expressions. Continued research and development in this area are crucial to further harnessing the potential of AI-powered machine translation systems. When comparing the translations generated by traditional MT systems with their 2019 counterparts (At-tall, 2019), it was found that they largely maintained their previous outputs. However, the newly developed LLM systems, such as BAI, consistently delivered translations that were substantially more accurate. This disparity between the limitations identified in previous studies (e.g., Al-Kharabsheh & Yassin, 2017; At-tall, 2019) and the potential for LLMs suggested by this study, alongside more recent research (e.g., Boughorbel & Hawasly, 2023; Banimelhem & Amayreh, 2023), hints at progress in machine translation, particularly within LLM-driven approaches. While recent research (e.g., Banimelhem & Amayreh, 2023) found that LLM systems performed below average compared to established MT systems on standard Arabic translation, this study suggests that LLMs may produce more accurate renditions of non-standard Arabic sentences, especially colloquialisms.

The study revealed several advantages of BAI over established MT systems. Firstly, BAI consistently provided translations that were closer to the human translations in At-tall (2019) and the intended meaning of the colloquial Arabic expressions, often correcting inaccuracies present in the outputs of GT and BT. This highlights the system's improved accuracy in translating the Arabic language. Secondly, the translations produced by BAI demonstrated a higher level of readability as indicated by the punctuation marks in the output even if they were not present in the source texts. This suggests that BAI can adapt to the stylistic subtle distinctions commonly found in Arabic expressions. One notable advantage of BAI was its interpretive depth. The system accompanied its translations with explanatory comments and contextual insights, indicating a deeper understanding of the underlying meaning behind the phrases. This additional layer of comprehension is particularly valuable for complex linguistic expressions, enhancing the overall quality and usefulness of the translations. The observed accuracy and interpretive depth of BAI suggest that its ability to conduct research on phrases before translating plays a crucial role. This method, although more time-intensive compared to the instant translations provided by GT and BT, appears to yield more accurate and insightful results.

Overall, this research presents a promising outlook for the future of machine translation. The advancements demonstrated by BAI, particularly its ability to handle complex linguistic expressions with accuracy, highlight the potential of AI-driven approaches to revolutionize the field. These findings emphasize the importance of further research and development in AI-powered LLM machine translation systems, as they hold the key to unlocking more accurate and contextually aware translations in the future.

Further research on LLM and MT tools could explore testing how these tools can accurately translate the complexities of other languages, particularly when it comes to translating

colloquialisms, and complex linguistic structures. Additionally, larger, more comprehensive studies could be carried out to further test these systems in translating complex structures. Research on the ethical implications of using LLM-powered machine translation systems like BAI can also be carried out to ensure responsible use. Furthermore, potential applications of LLM beyond MT can be explored, such as adapting it to more extensive speech recognition translation or text-to-speech systems. Finally, a human-AI collaboration model can be further developed and refined to improve translation quality and accuracy. These questions represent just a few of the many avenues for future research on AI and the broader field of LLM translation. By exploring these questions and others like them, researchers can gain a deeper understanding of the potential and challenges of these systems and pave the way for new breakthroughs in this emerging field.

Conclusion

This research aimed to explore two critical aspects of machine translation: the progress made since 2019 and the potential of LLM tools like BAI. By revisiting MT translations from 2019 (At-tall, 2019) and comparing them to current advancements the study aimed to shed light on evolving capabilities and identify promising developments. The qualitative analysis revealed substantial advancements in the accuracy and fluency of BAIs translations compared to its established counterparts. This remarkable performance points towards the transformative potential of LLM, with BAI's research-driven approach and interpretive depth highlighting the key strengths of this emerging technology.

However, it is important to acknowledge the limitations of this study. Firstly, the small sample size used in this study may not fully represent the diverse range of colloquial expressions present in Arabic. Conducting future research with larger datasets would provide a more comprehensive understanding of LLM system performance. Secondly, the reliance on qualitative analysis, while valuable for capturing nuanced translations, may benefit from complementing it with quantitative metrics to ensure a more objective evaluation. Incorporating independent evaluators would also enhance the reliability of the assessments. Despite these limitations, this study provides valuable insights into the potential of LLMs, such as BAI, in improving Arabic-English translations. It serves as a starting point for further research endeavors that can address these limitations and delve deeper into the capabilities and limitations of LLM systems in handling complex linguistic structures. Future studies with larger and more diverse datasets will contribute to a more comprehensive understanding of the effectiveness and applicability of LLMs in machine translation. This research represents an important milestone in the field of colloquial Arabic translation. It highlights the potential of AI-powered tools while emphasizing the importance of responsible and inclusive development. As developments move forward, the prospect of achieving accurate and insightful machine translation is within reach.

Funding
This research is not funded.
Acknowledgments
Not applicable.
Conflicts of Interest
The author declares no conflict of interest.
Authenticity
This manuscript is an original work

Artificial Intelligence Statement:
During the preparation of this work the author used Bing AI Chat in order to conduct the study. After using this tool, the author reviewed the content as needed and takes full responsibility for the content of the publication.

About the Author:
Hamad Abdullah H Aldawsari is an Assistant Professor of Translation Studies at the College of Sciences and Humanities, Prince Sattam bin Abdulaziz University (PSAU), Wadi Addawasir, Saudi Arabia. He completed his MA and PhD at the University of Birmingham, UK. His research interests include Translator Studies & Linguistics. ORCID ID https://orcid.org/0009-0001-2041-9760.

References
Al-Btoush, M. (2014). English Loanwords in Colloquial Jordanian Arabic. *International Journal of Linguistics*, *6* (2), 109-119. https://doi.org/10.5296/ijl.v6i2.5086

Aldawsari, H. A. H. (2023). Comparing the Performance of Google Translate and SYSTRAN on Arabic Lexical Ambiguity. *Arab World English Journal for Translation & Literary Studies*, *7*(3) 19- 34. http://dx.doi.org/10.24093/awejtls/vol7no3.2

Ali, M. A. (2020). Quality and machine translation: An evaluation of online machine translation of English into Arabic texts. *Open Journal of Modern Linguistics*, *10*(05), 524–548. https://doi.org/10.4236/ojml.2020.105030

Al-Jarf, R. (2023). Pluralization of social media loan terminology in colloquial Arabic. *British Journal of Applied Linguistics*, *3*(2), 60-68. https://doi.org/10.32996/bjal.2023.3.2.6

Al-Kharabsheh, A., & Yassin, O. (2017). Translation of colloquialisms in the Arabic-into-English subtitled film, the dupes. *International Journal of Comparative Literature and Translation Studies*, *5*(3), 18. https://doi.org/10.7575/aiac.ijclts.v.5n.3p.18

Almahasees, Z. (2021). *Analysing English-Arabic Machine Translation: Google Translate, Microsoft Translator and Sakhr*. Routledge. https://doi.org/10.4324/9781003191018

Alnamer, A., & Alnamer, S. (2018). The Use of Loanwords in Emirati Arabic According to Speakers' Gender, Educational Level, and Age. *International Journal of Applied Linguistics & English Literature*, *7*(4), 158- 176. https://doi.org/10.7575/aiac.ijalel.v.7n.3p.158

Al-Saidat, E. (2011). English Loanwords in Jordanian Arabic: Gender and Number Assignment. *Language Forum*, *37*(1), 59-72

Ameur, M. S. H., Meziane, F., & Guessoum, A. (2020). Arabic machine translation: A survey of the latest trends and challenges. *Computer Science Review*, *38*, 100305. https://doi.org/10.1016/j.cosrev.2020.100305

At-tall, S. M. (2019). *Comparative study between Google Translator and human translator in rendering colloquial Arabic expressions in the late Prime Minister Wasfi At-Tall's speeches into English*, (Unpublished Master's thesis). Yarmouk University, Irbid, Jordan.

Banimelhem, O., & Amayreh, W. (2023). Is ChatGPT a Good English to Arabic Machine Translation Tool?. In *2023 14th International Conference on Information and Communication Systems* (ICICS) (pp. 1-6). IEEE. https://doi.org/10.1109/ICICS60529.2023.10330525.

Boughorbel, S., & Hawasly, M. (2023). Analyzing Multilingual Competency of LLMs in Multi-Turn Instruction Following: A Case Study of Arabic. *SIGARAB Arabic NLP*. https://doi.org/10.48550/arXiv.2310.14819

Habash, N., & Diab, M. T. (2012). Arabic natural language processing: Why it matters and how to get started. In *Proceedings of the ACL-HLT 2012 Workshop on Computational Approaches to Arabic Script Languages* (pp. 1-9). Association for Computational Linguistics.

Harrat, S., Meftouh, K., & Smaili, K. (2019). Machine translation for Arabic dialects (survey). *Information Processing & Management*, *56*(2), 262-273. https://doi.org/10.1016/j.ipm.2017.08.003.

Huang, H., Wu, S., Liang, X., Wang, B., & Zhao, T. (2023). Towards Making the Most of LLM for Translation Quality Estimation. In *CCF International Conference on Natural Language Processing and Chinese Computing* (pp. 375-386). Cham: Springer Nature Switzerland. https://doi.org/10.3115/1073083.1073135

Hutchins, W. J. (2001). *Early Years in Machine Translation*. John Benjamins Publishing Company. https://doi.org/10.1075/sihols.97

Hutchins, W., & Somers, H. (1992). *An introduction to machine translation*. London: Academic Press.

Rinsum, H. (2023). CHATGPT, Bing Chat or bard? *Agrarzeitung*, *78*(39), 10–10. https://doi.org/10.51202/1869-9707-2023-39-010

Jibreel, I. (2023). Online machine translation efficiency in translating fixed expressions between English and Arabic (proverbs as a case-in-point). *Theory and Practice in Language Studies*, *13*(5), 1148–1158. https://doi.org/10.17507/tpls.1305.07

Joyce, M. (2008). *Anglo-American support for Jordan: The career of King Hussein*. New York: Palgrave Macmillan. https://doi.org/10.1057/9780230611641

Koehn, P. (2010). *Statistical machine translation*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511815829

Lee, T. (2023). Artificial intelligence and posthumanist translation: ChatGPT versus the translator. *Applied Linguistics Review*, *14*(4), 60-66. https://doi.org/10.1515/applirev-2023-0122

Morris, B. (2001). *Righteous victims: A history of the Zionist-Arab conflict*, 1881-2001. New York: Vintage https://doi 10.7440/histcrit21.2001.12

Okpor, M. D. (2014). Machine translation approaches: Issues and challenges. *International Journal of Computer Science Issues (IJCSI)*, *11*(5), 159. https://doi.org/10.20943/01201702.5457

Susser, A. (1994). *On both banks of the Jordan: A political biography of Wasfi al-Tall*. London: Frank Cass

Zakraoui, J., Saleh, M., Al-Maadeed, S., & Alja'am, J. M. (2021). Arabic Machine Translation: A Survey with Challenges and Future Directions. *IEEE Access*, *9*, 161445-161468. https://doi.org/10.1109/ACCESS.2021.3132488

Appendices

Appendix A

Arabic Test Suit

Table 1. *Arabic sentences in the test suite*

1. اللي ما بقبل البلد بضيقها البلد ما بتقبله بعزها
2. ما دام هالسراج فيه زيت خليه ضاوي
3. بدكو وطن بدون ما نضحي من أجله
4. القمح قمحي والبلاد بلادي والدم أردني
5. بجوز أنا أكون ضحية وبجوز غيري يصير ضحية
6. هل الأردن من أحسن الناس
7. ما حدا بموت ناقص عمر
8. اللي خبزه مش من قمحه عاش مذلول
9. هذا البلد أرضه وإحنا أهله
10. عقليات سايكس بيكو إحنا ضحية إلها

Note 1. Adopted from At-tall (2019, p. 15)

Appendix B

Google Translate Output

Table 2. *English translations by Google Translate*

1. What the country accepts with its distress, what the country accepts with its pride
2. As long as this lamp has oil, it will be bright
3. You want a homeland without what we sacrifice for
4. Wheat is my wheat, the country is my country, and the blood is Jordanian
5. I may be a victim and someone else may become a victim
6. The people of Jordan are among the best people
7. No one dies without age
8. He whose bread was not made from wheat lived humiliated
9. This country is its land and we are its people
10.Sykes-Picot mentalities: We are the victim of a god

Appendix C

Bing Translator Output

Table 3. *English translations by Bing Translator*

1. Who doesn't accept the country with its narrowness the country doesn't accept it with its dignity
2. As long as this lamp has oil in it, let it be lit
3. You want a homeland without sacrificing for it
4. Wheat is wheat and the country is my country and the blood is Jordanian
5. Maybe I will be a victim and maybe someone else will become a victim

ISSN: 2229-9327

6. The people of Jordan are among the best people
7. No one dies minus age
8. Whoever his bread is not from wheat lived humiliated
9. This country is his land and we are his family
10. Sykes-Picot mentalities we are a victim of it

Appendix D

Bing AI Chat Output

Table 4. *English translations and commentary by Bing AI Chat*

1. Whoever does not accept the country in its hardship, the country does not accept him in its glory. This is a proverb that expresses the idea of loyalty and patriotism. It implies that one should not abandon their homeland when it faces difficulties, but rather stay and contribute to its development and prosperity. Otherwise, they do not deserve to enjoy its benefits and honors when it prospers.
2. As long as this lamp has oil, let it shine. This is another proverb that expresses the idea of perseverance and optimism. It implies that one should not give up on their goals or dreams, but rather keep working hard and making the best of their resources until they achieve them. It also suggests that one should not waste their talents or potential, but rather use them to benefit themselves and others. This is a rhetorical question that expresses the idea of patriotism and sacrifice. It implies that one cannot expect to have a homeland without making some sacrifices for its sake, such as fighting for its freedom, defending its sovereignty, or contributing to its development. It also suggests that one should be grateful and proud of their homeland, and not take it for granted.
3. You want a homeland without sacrificing for it. This is a rhetorical question that expresses the idea of patriotism and sacrifice. It implies that one cannot expect to have a homeland without making some sacrifices for its sake, such as fighting for its freedom, defending its sovereignty, or contributing to its development. It also suggests that one should be grateful and proud of their homeland, and not take it for granted.
4. The wheat is my wheat, the land is my land, and the blood is Jordanian. This is a famous slogan from the Jordanian national movement in the 1950s that expresses the pride and patriotism of the Jordanian people and reflects the importance of agriculture and blood ties in the Jordanian culture and identity.
5. Maybe I am a victim and maybe someone else becomes a victim. This is a sentence that expresses the idea of uncertainty and vulnerability. It implies that one does not know what will happen to them or others in the future, and that they may face some harm or injustice. It also suggests that one is aware of

ISSN: 2229-9327

the possibility of being a victim or witnessing someone else being a victim, and that they are not indifferent to it.
6. The people of Jordan are among the best people. This is a sentence that expresses the idea of praise and admiration. It implies that the speaker has a high regard for the people of Jordan, and considers them to be good, noble, or virtuous. It also suggests that the speaker has some connection or affinity with the people of Jordan, or has experienced their hospitality or generosity.
7. No one dies before their time. This is a proverb that expresses the idea of fate and destiny. It implies that one's lifespan is predetermined by God or a higher power, and that nothing can change it. It also suggests that one should not fear death, but rather accept it as inevitable and part of the divine plan.
8. Whoever's bread is not from his wheat, lived humiliated. This is a proverb that expresses the idea of self-reliance and dignity. It implies that one should not depend on others for their livelihood, but rather work hard and earn their own bread. It also suggests that one who lives off the charity or generosity of others, loses their respect and honor.
9. This country is its land and we are its people. This is a sentence that expresses the idea of belonging and identity. It implies that the speaker has a strong connection and attachment to their country, and considers it as their home and source of pride. It also suggests that the speaker is loyal and committed to their country, and willing to protect and serve it.
10. We are the victims of the Sykes-Picot mentality. This is a sentence that expresses the idea of historical injustice and resentment. It implies that the speaker blames the Sykes-Picot Agreement of 1916, which divided the Middle East into spheres of influence for Britain and France, for the current problems and conflicts in the region. It also suggests that the speaker feels oppressed and exploited by the colonial powers and their successors.

ISSN: 2229-9327