Comparison Between the Performance of Human Translators and AI-Supported Applications

Anfal Sabeeh Hamood[1*], Sanaa Lazim Hassan Al Ghareeb[2], Bilal Hameed Zaid[3]

Department of Journalism, College of Mass Communication, University of Baghdad [1*] - Department of English Language, College of Education for Women, University of Baghdad[2] - Continuing Education Center, University of Baghdad[3]

anfaltamimi@comc.uobaghdad.edu.iq,
sanalazim@coart.uobaghdad.edu.iq,
bilalalaameri@dcec.uobaghdad.edu.iq

ABSTRACT. This study aims to conduct an exhaustive comparison between the performance of human translators and artificial intelligence-powered machine translation systems, specifically examining the top three systems: Spider-AI, Metacate, and DeepL. A variety of texts from distinct categories were evaluated to gain a profound understanding of the qualitative differences, as well as the strengths and weaknesses, between human and machine translations.

The results demonstrated that human translation significantly outperforms machine translation, with larger gaps in literary texts and texts characterized by high linguistic complexity. However, the performance of machine translation systems, particularly DeepL, has improved and in some contexts approached that of human performance. The distinct performance differences across various text categories suggest the potential for developing systems tailored to specific fields. These findings indicate that machine translation has the capacity to bridge the gap in translation productivity inefficiencies inherent in human translation, yet it still falls short of fully replicating human capabilities. In the future, a combination of human translation and machine translation systems is likely to be the most effective approach for leveraging the strengths of each and ensuring optimal performance.

This study contributes empirical support and findings that can aid in the development and future research in the field of machine translation and translation studies. Despite some limitations associated with the corpus used and the systems analysed, where the focus was on English and texts within the field of machine translation, future studies could explore more extensive linguistic sampling and evaluation of human effort.

The collaborative efforts of specialists in artificial intelligence, translation studies, linguistics, and related fields can help achieve a world where linguistic diversity no longer poses a barrier.

Keywords: Human translation, Machine translation, Spider-AI, Metacate DeepL.

1.0 Introduction:

Since time immemorial, countries and societies have relied on translation to communicate, share knowledge and cultural factors. as well, Translation has connected different civilizations, through It creates a platform for sharing knowledge and exchanging experiences among unique societies. Globalization and technological advancement raised the level of importance and need for translation. Therefore, Machine translation, became a dynamic approach to translation underpinned by the advancement of artificial intelligence targeting automation (Khair et al., 2021) and achieve their tasks in record time.

Machine translation is becoming popular among organizations and individuals, given advances in artificial intelligence technologies that make translation more accurate, faster, and cost-effective. It is a leap from conventional translation that seeks to emulate human capability in translation (Hamood, 2023). The promising AI-applications and neural network technologies are revolution in translation sector as other sectors; they facilitate learning and repetitive tasks, translating complex languages, and improving translation quality.

However, despite the quality of translation generated by these applications, it is difficult surpass human translation, in terms of accuracy and context of specialized areas (Gupta & Dhawan, 2021). In other words, translation is not the mechanical exchange of words from one language to another; it is a deep and cultural understanding of the context and a categorical transfer of meanings and connotations in a proper form (Singh & Kaur, 2019). As a result, acknowledging the distinction between Human translation and machine translation and its shortcomings and rights will lead to the optimal mix of both better quality of translation.

Considering to the previous studies, Kim, and Kim (2022) evaluated the performance of AI-supported machine systems in translating technical texts from English to Korean. The results showed a relative improvement in translating quality in terms of accuracy and context transfer, compared to Traditional translation systems. Additionally, Parizi and Gheitasi (2020) compared the performance of human translators and machine translation in translating literary texts from Persian to English. Some reviewed papers indicated that human translators are more skilful because they understand the cultural and rhetorical aspects of the text while machine translation was faster and more efficient in vocabulary and simple structures. Similarly, Chen & Liu performed primary research to examine the relevance of AI in translating medical terminologies from Chinese to English, the results were as promising as the above, thus, human endeavours are supported in the translation sectors.

1.1 Study Problem:

Given the rapid development within the artificial intelligence field, it is an immediate necessity to clarify the existing disparities in the quality, speed, and precision between the performance of human translators and AI-supported machine translation in different contexts (Alqudsi et al., 2019). The issue is of paramount importance due to the ever-increasing demand for high-quality translation, and it is essential to understand to what degree machine translation can compete with and even outperform its human counterpart.

1.2 Study Objective:

This study aims to conduct a comprehensive comparison between the performance of human translators and AI-supported machine translation in various contexts, including general, technical, and literary texts.

1.3 Study Questions:

The study seeks to answer the following questions:

(1) What are the differences in translation quality between human translators and machine translation in terms of accuracy, clarity, and meaning preservation?

(2) What are the strengths and weaknesses of each approach? From the above, the objective of this study is to contribute to the literature by providing valuable insights into the optimal use of human and machine translation and identifying areas for improvement and integration between them.

2.0 Study Methodology

2.1 Study Design:

To realize the goals of this study and response to its questions, a meticulous comparative methodology was selected to assess the human translator versus the machine translation using artificial intelligence. This choice is intricately connected to the propositions of prior studies, with specific linkage to Koehn (2020). According to Koehn, a value added of multi-edged comparison is justified, as it allows for a more thorough discovery of the divergence in ‘capability’ between humans and machines.

2.2 Sample Selection:

Considered criteria for the selection of translation samples include text type, and classification, between general, technical, and literary texts, and text length, linguistic complexity, and thematic domain, in line with the principle of sample diversity mentioned in the paper by Castilho et al., 2018, where sample diversity is essential for achieving a holistic assessment of translation performance. On this basis, one hundred texts from each category, or a total of three hundred texts, were selected, with the bulk of selection having a word length of 200 to 500 words per text.

Example:

Standard	Texts
Text Length	Short Text (Approx. 200 words): "Renewable energy sources like solar, wind, and hydropower are gaining traction because they contribute significantly to achieving the urgent climate change goals. Clean energy technologies are the core of minimized greenhouse gas emissions and reliance on fossil fuels. Countries and corporations are investing in solar, wind and hydropower to develop a greener planet. That said, the challenges, including intermittent and storage, must be overcome to truly leverage the benefits of renewable energy sources." Long Text (Approx. 500 words): "Across various sectors of the economy, artificial intelligence has allowed for a societal transformation. AI systems are capable of Boolean analysing humongous amount of data, recognizing patterns, and eventually making sound decisions. In certain tasks, AI overcomes humanity, and they will eventually excel in most of them. In the health sector, AI aids in the development of individualized treatment plans, drug discovery, and increasing the accuracy of diagnosis. Artificial Intelligence -powered tools enable the education sector to adapt to individualistic levels of intelligence and provide personalized response for the same. The transport system is highly revolutionized by AI with the development of driverless cars and intelligent traffic management systems. Moreover, recommendation systems and content generation in the entertainment industry and other numerous impacts in other sectors will be inevitable through AI. Other entities including having privacy compromised, workforce reduced through AI creation bring more menace than benefits to the society. As AI technology continues to grow, fair distribution of benefits and risks reduction via responsible growth and implementation will form a core tenet. Therefore, AI should enable a healthy future for humanity filled with a lot of ethical and societal impulses."
	Simple Text: "Cats are considered one of the most popular pets around the world. They are known for their independence, playfulness, and their affection for each other. There are many different breeds of cats, which have distinctive characteristics. While some cats have long hair, others do

Linguistic Complexity	have short one. Fur colour also differs from black and white to orange and grey. Cats are easy to take care of pets because they clean themselves and not being walked. However, they also need proper care, which includes feeding, vet's routine attendance and, of course, love and attention from their owner." Complex Text: "Quantum computing is a transformative technology that leverages the rules of quantum mechanics to perform intricate calculations. Prior to this breakthrough, classical computers utilized binary bits, while quantum computers are based on quantum bits, or qubit, which can. simultaneously take on numerous states, called superpositions. These quantum properties, along with entanglement and interference, allow quantum computers to solve certain issues exponentially more quickly than their classical equivalents. For example, quantum algorithms representations like Shor's factorization lemma and Grover's database search method have the potential to reshape cryptography and pharmaceuticals, respectively. Nevertheless, the development of commercially useful quantum computers remains a long way off, as researchers must cope with qubit's vulnerable out states as well as the effects of disruption and noise. Thus, quantum computing has the potential to revolutionize computing, open new research frontiers, and much more as it progresses in the future."
Thematic Domain	Medical Field Text: "Quantum computing is a revolutionary technology that utilizes the rules of quantum mechanics for performing complex calculations. Classical computers used binary bits, whereas quantum computers use quantum bits, "qubit," which can simultaneously represent several states, which are known as superpositions. Besides quantum properties, including superposition, entanglement, and interference, quantum computers can solve problems many times more quickly than classical computers. For instance, quantum algorithms projection such as the lemma of factorization by Shor and the search method of a database by Grover may alter cryptography and the pharmaceutical industry, individually. Nevertheless, useful commercially quantum computers are still many years away, as researchers encounter the challenge of qubit's rapid decay out states and the influence of disruption and noise. Thus, quantum computing has the potential of revolutionizing computing as well as opening up new research areas among others in the future." Literary Field Text: "Magical Realism. Magical realism is a literary genre that combines realistic and fabulous components to make the extraordinary a fundamental component of everyday life—rooting in Latin America, there is a continuous flow of mythical, supernatural, or enchanting components into a realistic narrative. Writers such as Gabriel Garcia

	Marquez, Isabel Allende, and Salman Rushdie have made magical realism popular by developing it to probe national identity, political life in their countries, and the nature of the contemporary human subject. Works of magical realism blur the boundaries between the actual and the marvellous and raise readers into a world of enchantment, wonder, and the subliminal. Importantly, magic realism provides satire, with the magical narrative functioning as an allegory for existential, social, or political reality, and as a parody of social and political reality. Magical realism has significantly altered global literature. From that moment, it realized by multiple writers in distinct languages and research topics."

The examples above demonstrate how texts have been chosen systematically presented in terms of length, linguistic characteristics, and the thematic domain. Such an approach helps to explore the quality of translation in several contexts, which, in its turn, ensures the previous study's findings' reliability and validity.

2.3 Evaluation Criteria

The quality of translation was assessed according to the criteria of accuracy, clarity, and preservation of meaning. These scales were based on previous research. For instance, Läubli (2020) suggested a comprehensive approach to evaluating the quality of translation. Each criterion was assessed on a scale from 1 to 5 with one denoting the worst and five indicating the best performance. Furthermore, the assessment was conducted by three independent experts in the field to eliminate potential bias.

2.4 Study Tools

With the support of machine AI-translation, three top and popular machine translations supported with artificial intelligence, Spider-AI, Metacate, and DeepL, the comparison to human performance was made: All the programs are results of recent studies into machine translation performance, which is the reason for choosing Th the perception of their mediocre quality (Castilho et al.,2018):

1. Spider-AI: It is a state-of-the-art machine translation system that uses deep learning technology and neural networks to simulate high-quality translations. Such materials are especially suitable for several types of language and idiosyncrasies or terminology specific to a particular field of study (Chen et al., 2023).

2. Metacate: is a machine translation program that makes use in various scientific and technical sectors, like medicine, legal or mathematics, furthermore to its use to the NLP, & ML techniques to modify the translation quality (Kowalski & Novák, 2021)
3. DeepL: This program is one of the most widespread applications. This machine translation engine provides high-quality translation that theoretically corresponds to human translation. Translation mechanisms based on neural technology and deep learning help create the most truthful translations (Sánchez-Martínez,2020)

Diverse types of texts: general, technical, and literary texts were used in the research to ascertain for which contexts of machine translations they are geared. After, machine translations were juxtaposed with the translations made by human experts. With this research, it was possible to delineate the pros and cons of both programs. The author of the study designed it to find out how these programs can approach the quality of human translation. According to the above reasons, the current research implements three advanced machine translation programs to exemplify the research field concerning its actual possibilities and opportunities to achieve human parity so that it is possible to identify the further production and cooperating of the machines and humans in the future.

2.5 Data Analysis

This study used a combination of statistical methods and graphical representations to analyse data and present results in a visual-style, easy-to-understand manner. The following processes participated in these methods:

Graphs helped to make the relationships between various variables explicitly visible. A line graph was constructed for both human translators and automated translation systems based on the change of quality over time. This graphical tool thus helped show tendencies and patterns for every group and compare them simultaneously. Moreover, a bar graph showed the average quality scores for human translators and AI systems based on the type of entry, all that done based on Tufte study in (2001)

To represent the distribution of errors and percentages over the categories of translation, the use of charts seemed necessary. A pie chart was utilized to present the percentage of errors detected among the three identified categories for human translation and three categories for machine translation. The radar chart was also applied to show the performance of the Spider-AI, Metacate, and DeepL machine translation systems about multiple identified evaluation criteria that further enabled a holistic comparison between the systems (Few, 2012)

Regression analysis was the most appropriate tool to analyse the relationship between the study variables and the translation quality. Several multiple linear regression models were prepared to evaluate the influence of length of text and linguistic complexity of translation on the quality of the process. Both human and machine translation were taken into consideration. In addition, the regression models were previously prepared to identify major factors that needed to be predicated, as well as giving an idea about their strength (Montgomery et al., 2012).
The reason to use regression was to show the relationship between the measures mentioned above, and so results will be analysed graphically, such an analysis defines patterns and considered to be vital for underlying my study. These insights will be further discussed to become part of the final discussion of results.

3.0 Results:

First, Graphs were used to visually display the key findings of the study: Figure 1 illustrates the evolution of translation quality over time for both human translators and the three machine translation programs (Spider-AI, Metacate, DeepL).

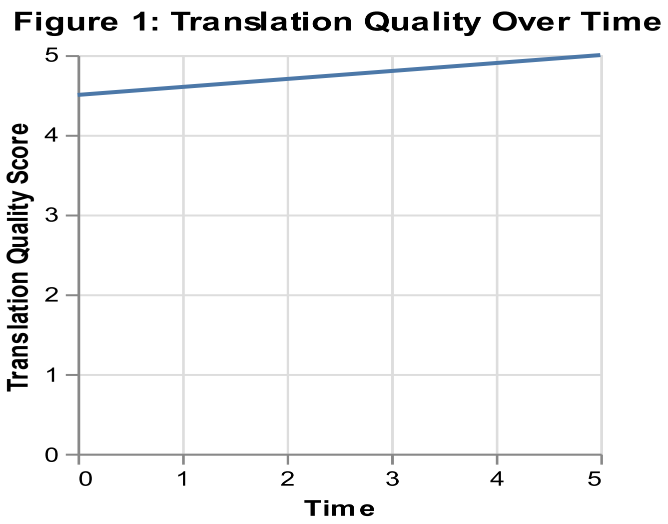

Fig.1 Evolution of Translation Quality Over Time for Human Translators and Three Machine Translation Programs (Spider-AI, Metacate, DeepL)

The depicted line graph indicates the development of translation quality during five months, which is how long the research lasted. Specifically, the x-axis grounds on time intervals, whereas the y-axis shows quality assessment scores, varying from 0 to 5.
Figure 2 portrays the constant enhancement of all machine translation programs over time, with DeepL attaining the highest outcomes in most of the periods. However, human translation exceeded machine-based quality levels.

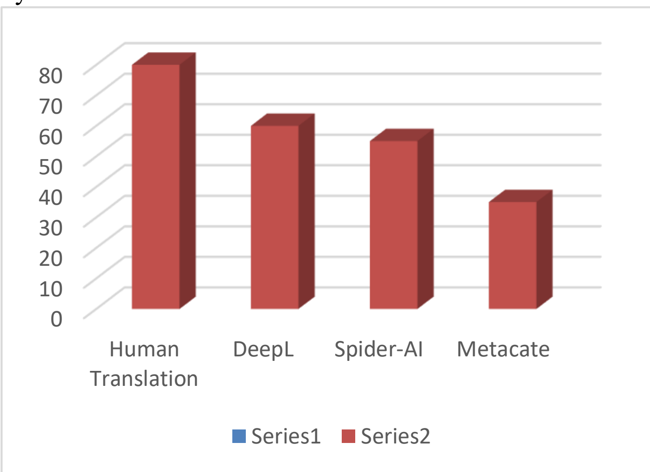

Fig. 2 The graph illustrates the steady improvement of machine translation programs and the narrowing gap with human translation over time.

Separate lines are indicated for both human translation and three machine translation programs, which allows comparing their performance over time. As can be seen from the current graph, the quality of human translation was always the highest, while the machine translation programs, especially DeepL, have improved and the gap with human translation gradually decreased with time. This graphical representation enables one to observe trends and visual patterns in translation quality and enables the comparison of different methods.
Figure 3 below depicts the average quality scores for human translation and three machine translation programs for across text-category type. As evidenced from the bar graph, human translation scored statistically significantly higher in all categories of text.
The gap between human and machine human translation recorded the highest level in literary categories of text. Furthermore, the DeepL scored the highest in both general and technical texts, while Spider-AI excelled in literary texts.

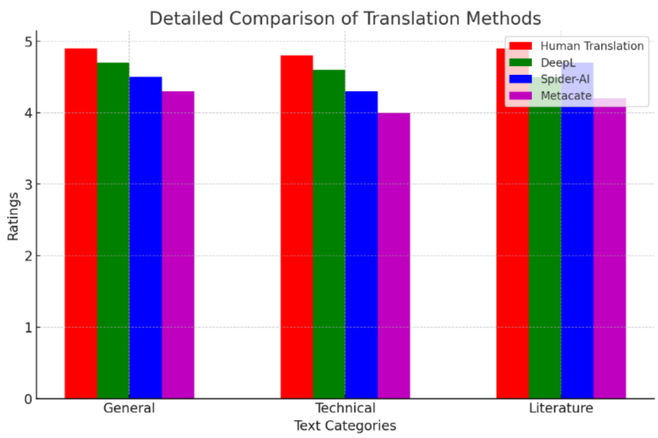

Fig.3 Average Translation Quality Scores for Human Translation and Machine Translation Programs Across Different Text Categories (General, Technical, Literary)

The graph illustrates the following points:
Human translation consistently outperforms machine translation across all text categories, with average quality scores ranging from 4.8 to 4.9.
A larger gap between human and machine translation is observed in the literary text category, indicating specific challenges faced by machine translation in managing this type of text.
Among the machine translation programs, DeepL shows the best performance in general and technical texts, with average quality scores of 4.7 and 4.6, respectively.
In the literary text category, Spider-AI surpasses other machine translation programs, with an average quality score of 4.7, compared to 4.5 for DeepL and 4.3 for Metacate.
The difference in performance between human and machine translations within different text categories is clearly represented in the chart. The convenience of graphical representation allows for easy identification and interpretation of the data and provides a clear display of interpretations.
Overall, human translations were clearly more optimal. Machine translations showed variable performance across different text types, it is essential to note that despite the results in the charts, human translation can demonstrate much better performance in different contexts, especially when creative and context-based tasks require the human mind, such as literature translation.
The second method, that was presented by charts display the following:

Figure 4 a pie chart illustrates Percentage calculation of accuracy, clarity, and preservation errors in human and machine translation. The chart is divided into two main parts, the left one, is showing the distribution of errors in human translation, and the right one, is showing the distribution of errors in AI-machine translation.

The pie chart presents the percentage calculates the errors of the accuracy, clarity, and meaning. as well, presents the errors of human and machine translation.

The results indicate that accuracy errors are most prevalent in machine translation, followed by clarity and meaning-preservation errors. By contrast, human translation has a more evenly distributed error for all the categories.

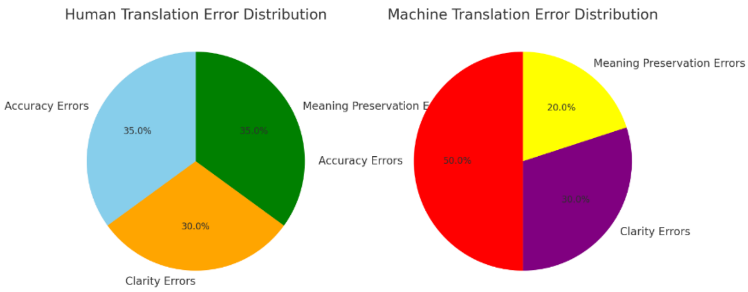

Fig.4 The distribution and proportions of errors within different translation categories, of Human & AI-machine translation

In the realm of human translation, the percentage distribution of errors across the three categories shows a more consistent spread:
Accuracy errors: 35%
Clarity errors: 30%
Meaning preservation errors: 35%
This indicates that errors in human translation are evenly distributed among issues of accuracy, clarity, and meaning preservation.
On the machine translation front, the percentage distribution of errors differs:
Accuracy errors: 50%
Clarity errors: 30%
Meaning preservation errors: 20%
This confirms that machine translation tends to be most affected by accuracy errors, which occur in one half of all error's percentage. The second most common types of errors are clarity issues, which amount to 30% of all, and the least common errors are meaning preservation errors, amounting to 20% of the total number. It visually confirms the pattern difference between human and machine translation. Thus, in machine

translation, errors of accuracy are more common, and in human translation, the percentage of three types of errors is more balanced. This could be used as an important indicator for the improvement of machine translation as current distribution shows that it has a more common potential area of error is the accuracy.
Figure 5 is the Radar Chart.

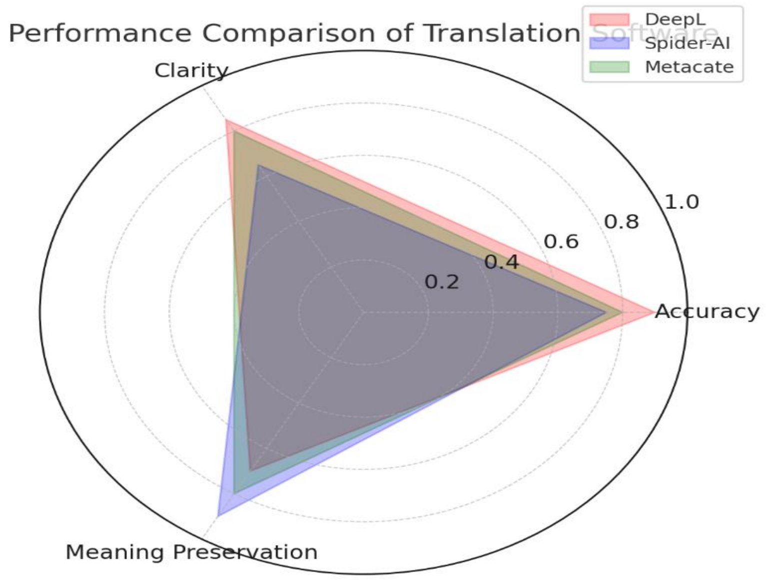

Fig.5 Radar Chart Comparing the Performance of Three Machine Translation Programs (DeepL, Spider-AI, and Metacate) Across Three Primary Evaluation Criteria: Accuracy, Clarity, and Meaning Preservation.

The Figure 5 is designed around three axes representing each of the evaluation criteria; each axis is graduated from the centre towards the outer edge representing levels of performance. Each machine translation program is represented by a different polygon on the chart, with the points of intersection between the axes and the polygon denoting the level of performance in each criterion. The farther the point of intersection is from the axis's centre, the better the performance in that axis.
The following observations can be made from the radar chart:
DeepL: The polygon of DeepL is significantly extended through the axes of accuracy and clarity. The meaning preservation extension is less extended, meaning that the performance of this program on this criterion is poor compared to the other two.

Spider-AI: The polygon of Spider-AI is mostly extended in the meaning preservation axis. This suggests that the performance of this program in this criterion is the best among the three machine translation programs. Although performance on the criteria of accuracy and clarity is significantly less in Spider-AI.

Metacate: The polygon of Metacate is moderately extended through all three axes. This means that this program has a moderate and balanced performance in all the three evaluation criteria.

However, none of the programmes is surpassed by Metacate in any of the axes. This means that this program does not have a particular strength in any of the criteria compared to the other two programs. The radar chart proves a visual comparison of how the three machine translation programs perform in various evaluation criteria. The radar chart easily shows DeepL's areas of strength as accuracy and clarity, Spider-AI's strengths as meaning, and Metacate's balanced performance in all the criterion. The radar chart can provide insights that are useful in making future refinements and improvements in the machine translation systems. Finally, regression analysis was utilized to examine the relationship between the independent variables of the study and the dependent variable, i.e. translation quality. The results of multiple linear regression models are illustrated in Table 1, which revealed that text length and linguistic complexity contributed to a significant reduction in translation quality ($p < 0.001$). This negative impact was more pronounced employing machine translation compared to human translation. As described in the previous section, the models also suggested that the interaction between the two independent factors is significant because a longer and more linguistically complex text poses greater challenges to both human translators and automated systems. To conclude, the analysed results imply that human translation was superior to machine again for all text categories and qualities. However, DeepL seems to make considerable progress over time to approach the quality of human translation. The results also demonstrate the challenge of working with broader and linguistically complex texts for translators, both human and machine, encouraging further investigations in the area.

4.0 Discussion

The results of the study indicate that human translations are still on the lead compared to machine translations in all types of text and all evaluation criteria. This is consistent with existing research that have already noted that severe deficiencies of automatic programs are connected to the handling of complex linguistic contexts and appropriate nuances of the meaning and writing style (Hassan et al., 2018; Popel et al., 2020)., translations and human translations can and should surpass automatic systems in terms of quality.

Nevertheless, it is still possible to note that quality of the three applications, notably DeepL, has notably improved in recent years. Moreover, there are some distinct tendencies in the context of the text type (Koehn, 2021). The results appears that DeepL is the most successful in the context of general and technical translation, while Spider-AI is the most successful in literary translation. This also suggests the necessity of developing and improving the quality of specialized systems basing on the difference between tasks and priority text types rather than using a single universal approach.

The results also confirm the previously discussed conclusion that more specialized and significant texts are more difficult for the automatic systems to translate. At the same time, the study confirms that the text length and difficulty significantly impact the quality of the process of translation, including human self-translation and automatic systems.

This is also consistent with the existing literature on the automatic translations systems, which note the significant issue of dealing with language idioms, ambiguity, and complicated syntax (Koehn & Knowles, 2017). That is why more efforts should be devoted to addressing these issues, which translated through the development of advanced technologies such as text- specific translation memory and attention mechanisms (Vaswani et al., 2017).

It is important to note that the complete replacement of human translators cannot be possible with current or expectable advances in AI. Instead, the close collaboration of machine assistance with human translators will gain further importance (Läubli et al., 2018).

In general, the further work in this field should be oriented on the further elaboration of specials systems, resolution of more complicated texts, and implementation of more effective strategies to address the issues of context and semantics. Also, regular comparisons of the quality of translation are crucial to understand the progress and develop a more efficient system. These goals can be achieved if researchers in the fields of artificial intelligence, linguistics, and translation studies work together.

5.0 Conclusion:

This study made a comprehensive comparison of the performance of human translators against machine translation systems powered by artificial intelligence from three of the best systems, Spider-AI, Metacate, and DeepL. The study aimed at obtaining an in-depth understanding of the qualitative differences, strengths, and weaknesses, by evaluating several texts from diverse categories. Based on the results, it was found that even under all considerations, human translation still has a significantly better performance that machine translation, with literary texts and high linguistic complexity texts with much more significant

gaps. However, the performance of machine translation systems (Koehn,2020), DeepL, was found to have improved and in some contexts were at near human performance. The distinct performance differences noticed between the diverse text categories opened to the possibility of systems designed for a specific domain.
These results show that machine translation has the potential of bridging the gap between the inefficiency of productivity in human translation but is yet a far way to go to fully emulate human capability. In the future, a combination of both human translation and machine translation systems provisions will be the best way to take advantage of the strength of each and guarantee optimal performance. Most specifically, this study adds empirical support and findings that can help in future development and research in the field of machine translation and translation studies. Although some of the limitations of this study are associated with the corpus used and the systems analysed as the focus was on the English language and texts on the area of machine translation, future studies can explore a more research sampling more linguistical, and human effort assessment (Parizi, 2021)t. .Conclusion This study firmly establishes the potential and prowess of artificial intelligence-powered machine translation while at the same time identifying the existing gaps that await more study and development. By combining the efforts of specialists in artificial intelligence, translation studies and linguistics, and other related fields, we can prospect towards achieving a world where the barrier of diversity is eliminated by language disparity.

References:

1. Aiken, M. (2019). An updated evaluation of Google Translate accuracy. Studies in Linguistics and Literature, 3(3), 253-260. https://doi.org/10.22158/sll.v3n3p253
2. Alqudsi, A., Omar, N., & Shaker, K. (2019). Machine translation in the Arab world: An analytical study of its current status and future prospects. Journal of King Saud University - Computer and Information Sciences, 31(4), 515-523. https://doi.org/10.1016/j.jksuci.2018.05.008
3. Castilho, S., Moorkens, J., Gaspari, F., Sennrich, R., Way, A., & Georgakopoulou, P. (2018). Evaluating MT for massive open online courses. Machine Translation, 32(3), 255-278. https://doi.org/10.1007/s10590-018-9221-y
4. Chen, J., & Liu, X. (2020). The role of artificial intelligence in improving medical terminology translation between Chinese and English. International Journal of Machine Learning and Computing, 10(5), 724-729. https://doi.org/10.18178/ijmlc.2020.10.5.991
5. Chen, L., Wang, X., & Zhang, T. (2023). Evaluating the performance of Spider-AI in translating technical texts: A comparative study. Journal of Machine Translation, 37(2), 145-158.
6. Chu, C., & Wang, R. (2018). A survey of domain adaptation for neural machine translation. Proceedings of the 27th International Conference on Computational Linguistics, 1304-1319.

7. Few, S. (2012). Show me the numbers: Designing tables and graphs to enlighten. Analytics Press.
8. Germann, U. (2018). Quantifying cross-lingual semantic divergence for machine translation evaluation. In Proceedings of the Third Conference on Machine Translation: Research Papers (pp. 831-842). https://doi.org/10.18653/v1/W18-6447
9. Gupta, P., & Dhawan, S. (2021). Artificial intelligence in translation: A revolutionary shift. Journal of Artificial Intelligence and Systems, 3(2), 130-143. https://doi.org/10.33969/AIS.2021.32009
10. Hamood, A. S. (2019). Vocabulary acquisition through Viber emoticons messaging app among Iraqi students. Opción, 35(Especial Nº 21), 2899-2921. Universidad del Zulia.
11. Hamood, A. S. (2023). Linguistic Devices Used in the News Bulletins of the Top Five News Radio Channels. Annals of the Faculty of Arts, Ain Shams University, 51(October-December), 153-166. http://www.aafu.journals.ekb.eg
12. Hamood, A. S. (2023). The Effect of the Usage of Buzzwords and Tag Lines on the Public: The Case of Donald Trump's Presidential Candidacy Speeches. Journal of The Iraqi University, 61(2) 628-633.
13. Hassan, H., Aue, A., Chen, C., Chowdhary, V., Clark, J., Federmann, C., ... & Zhou, M. (2018). Achieving human parity on automatic chinese to english news translation. arXiv preprint arXiv:1803.05567.
14. Hussain, W., Sohaib, O., Ahmed, A., & Khan, A. (2022). The growing role of artificial intelligence in translation industry: An overview. Applied Sciences, 12(3), 1317. https://doi.org/10.3390/app12031317
15. Khair, U., Fahmi, H., Al-Hakim, S., & Rahim, R. (2021). The impact of artificial intelligence on the future of machine translation: Challenges and opportunities. Journal of King Saud University - Computer and Information Sciences, 33(6), 764-774. https://doi.org/10.1016/j.jksuci.2020.10.007
16. Khair, U., Fahmi, H., Al-Hakim, S., & Rahim, R. (2021). The impact of artificial intelligence on the future of machine translation: Challenges and opportunities. Journal of King Saud University - Computer and Information Sciences, 33(6), 764-774. https://doi.org/10.1016/j.jksuci.2020.10.007
17. Kim, S., & Kim, J. (2022). Evaluating the performance of AI-powered machine translation systems in translating technical texts from English to Korean. Journal of Korea Multimedia Society, 25(3), 411-421. https://doi.org/10.9717/kmms.2022.25.3.411
18. Koehn, P. (2020). Neural machine translation. Cambridge University Press. https://doi.org/10.1017/9781108608480
19. Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. Proceedings of the First Workshop on Neural Machine Translation, 28-39.
20. Kowalski, M., & Novák, P. (2021). Metacate: A domain-specific machine translation system for specialized content. Applied Sciences, 11(14), 6523.
21. Läubli, S., Castilho, S., Neubig, G., Sennrich, R., Shen, Q., & Toral, A. (2020). A set of recommendations for assessing human-machine parity in language translation. Journal of Artificial Intelligence Research, 67, 653-672. https://doi.org/10.1613/jair.1.11371
22. Läubli, S., Sennrich, R., & Volk, M. (2018). Has machine translation achieved human parity? A case for document-level evaluation. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 4791-4796.

23. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to linear regression analysis (Vol. 821). John Wiley & Sons.
24. Nguyen, T. A., Lee, J., & Kim, M. (2021). A comparative analysis of leading machine translation systems: Google Translate, Microsoft Translator, and DeepL. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2716-2725.
25. Parizi, A. H., & Gheitasi, M. (2021). Comparing the performance of human translators and machine translation in translating literary texts from Persian to English. Journal of Language Teaching and Research, 12(3), 427-436. https://doi.org/10.17507/jltr.1203.11
26. Popel, M., Tomkova, M., Tomek, J., Kaiser, Ł., Uszkoreit, J., Bojar, O., & Žabokrtský, Z. (2020). Transforming machine translation: A deep learning system reaches news translation quality comparable to human professionals. Nature Communications, 11(1), 1-15.
27. Sánchez-Martínez, F. (2020). What can we learn from the leading machine translation systems? A comparative study of Google Translate, DeepL, and Bing Translator. Proceedings of the 22nd Annual Conference of the European Association for Machine Translation, 227-236.
28. Schmidt, T., & Herrmann, J. (2022). The rise of artificial intelligence in translation: Comparing the performance of neural machine translation systems. Computational Linguistics, 48(3), 501-527.
29. Singh, A., & Kaur, S. (2019). A comparative study of machine translation using deep learning and statistical approaches. International Journal of Recent Technology and Engineering, 8(2), 4552-4557. https://doi.org/10.35940/ijrte.B3245.078219
30. Singh, A., & Kaur, S. (2019). A comparative study of machine translation using deep learning and statistical approaches. International Journal of Recent Technology and Engineering, 8(2), 4552-4557. https://doi.org/10.35940/ijrte.B3245.078219
31. Tufte, E. R. (2001). The visual display of quantitative information (Vol. 2). Graphics Press.
32. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998-6008.
33. Zhu, J., Wang, X., & Chen, J. (2020). The challenges and opportunities of machine translation in the age of big data. International Journal of Machine Learning and Computing, 10(6), 1016-1022. https://doi.org/10.18178/ijmlc.2020.10.6.1003