A Review of Major Translation Technologies in Literature: Bridging to Hybrid Systems

Taruna Sharma[1], Nandini Sharma[2*], Arpit Dwivedi[3] and Daksh Thakur[4]

[1]Associate Professor, Applied Science Department, HMRITM, Hamidpur, New Delhi, 110036, Delhi, India

[2,3,4] Students, CSE Department, HMRITM, Hamidpur, New Delhi, 110036, Delhi, India

* Corresponding author

doi: https://doi.org/10.21467/proceedings.178.8

ABSTRACT

Literature has served as a means of expression for centuries. However, the earliest attempts to overcome language barriers can be traced back to the 1200 years old manuscript by Al Kindi on deciphering cryptographic messages. In the second millennium BC, the Sumerian poem Gilgamesh, was translated into Asian languages marking the earliest known literary translation. With the advent of computers in late mid twentieth century, translation technology began to take form. Translation involves conveying the essence of a text from one language into another. Through translation, literature gains the ability to transcend boundaries, enabling writers to communicate across time and cultures. Translated works contribute to the enrichment of the target language by introducing new terms and concepts, thereby benefiting the linguistic landscape. Eventually, along with the help of Machine Translation technology, it took the form of present-day existing techniques used in translation. The study provides a critical and comprehensive overview of the evolution of the said technology over the years. It familiarizes with the use of the translation in literature. It examines the challenges faced by each technique and how they were addressed by subsequent innovations. Furthermore, it will reflect on the different approaches taken in this field over the years: machine and human. Finally, the article will showcase the present-day challenges faced and speculate on the possibilities of its future advancement.

Keywords: literature, translation, neural machine translation, hybrid translation

1 Introduction

Throughout history, literature has served as a timeless mirror reflecting the complexities of human experience and the enduring quest for knowledge and meaning. However, each culture, enriched with its own set of rules, connotative and denotative notations for the language used, posed a challenge to traverse it across other cultures and generations. The transcendence of language barrier became feasible only with the 1200 years old manuscript on cryptographic messages. Thus, the need for translation emerged. Translation is commonly described as conveying the essence of a source-language text through an equivalent target-language rendition. However, this concise definition fails to meet the criteria of being 'both inclusive and exclusive' for a formal definition, but also because it presents translation as if it were a fixed and static concept of homogenous nature, which is shared by different cultures and nations at all times [1]. The notion of meaning and equivalence in translation varies across cultures and has encompassed different practices throughout historical periods. Nevertheless, there have been traces of the everchanging field of translation studies with techniques ranging from manual translation to use of NMT. From the earliest attempts at translation by scribes labouring over manuscripts to the cutting edge technologies of the digital age, the evolution of translation technology in literature is a testament to relentless pursuit of understanding and connection by the humanity. The following study will discuss about the different technologies evolved over course of time and how each of them resolved the challenges faced by its predecessor. Further, an analysis is presented on whether hybrid approaches combining machine translation

Proceedings DOI: 10.21467/proceedings.178; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-984081-8-1

technology with human intervention or post editing techniques enhance the quality and fidelity of translated literary texts or not.

2 Evolution of Technology

The translation technology has evolved over the course of time with significant advancements. It has been noted that each evolved technology built upon the achievements and addressed the limitations and challenges faced by its predecessor. The following is a brief history outlining the key stages of developments.

2.1 Manual Translation

Manual translation, regarded as the earliest form of translation, dates back to the second millennium BCE, when the Sumerian poem Gilgamesh was translated into Asian languages. This marked the first recorded effort to convey a text's meaning across linguistic boundaries. The translation in that era was primarily done by scribes, monks or bilingual individuals as shown in Figure 1. They would translate the text word for word by hand or paraphrase. They relied entirely on their linguistic expertise, cultural understanding, and interpretive skills. It was usually done in monasteries or scriptoria. A notable example is the Septuagint, the Greek translation of the Hebrew Bible, which required the collaborative effort of approximately 70 scholars to ensure its accuracy and consistency. This technique however was time consuming and solely dependent on human expertise. As it was done by hand, it had limited linguistic accuracy and a high risk rate of errors. Cultural nuances, idiomatic expressions and references were often lost or altered in the end result. Despite the setbacks, manual translation provided the only means to bridge linguistic divides for centuries, setting the stage for future innovations. These translations played a crucial role in preserving and transmitting ancient knowledge, allowing cultural and religious texts to survive across generations and geographical boundaries.

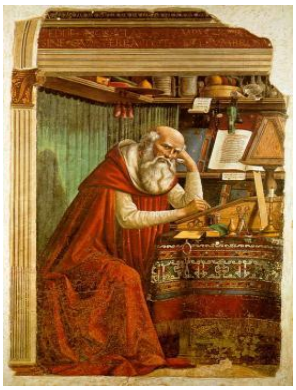

Figure 1. *St. Jerome in His Study* : A depiction of early manuscript translation [2]

2.2 Mechanical Translation

During several time periods of Middle Ages, same texts were translated from different manuscripts copies. Translation was done with the assistance of mechanical aides such as dictionaries, glossaries, and phrasebooks as illustrated in Figure 2. These mechanical aides were primarily used to enhance efficiency and provide consistency, particularly when translating texts or legal documents that required high accuracy. Notable figures like Roger Bacon, a 13th century scholar, proposed the concept of using mechanical devices for translation, as described in his "Opus Majus" (1267). While these ideas were theoretical and not practically realized, they hinted at the future potential of technology in linguistic endeavours. Translation during this period often involved multiple steps, including referencing existing manuscripts and consulting available mechanical aides. This method was reliant on manual labour and a lack of consistency was

observed. Automation was also limited in this technique, making the process both time-consuming and prone to human error. Despite these challenges, the use of mechanical aides paved the way for the development of more systematic and methodical translation processes. The early efforts contributed to the broader advancement of knowledge exchange between cultures. However, inconsistencies in translation were still a significant issue due to the variability in the quality of manuscripts and the individual expertise of translators. As a result, translation work in this period required a high level of human intervention making it a labour intensive task. The Middle Ages were thus a formative period, evolving the groundwork for future innovations in translation technology.

Figure 2. Dante Alighieri depicting the poet's role in translation using aid of books and references [3]

2.3 Early Machine Translation

With the start of Industrial Revolution, early machine translation systems, including rigid rule based and dictionary based approaches, emerged. This marked the advent of computational approaches to translation, laying the foundation for the rapid development of these techniques over time. While early systems had limited practical utility, they played a crucial role in shaping the trajectory of machine translation research and development, paving the way for future innovations in the field. These early systems were often restricted by the linguistic complexity of the languages involved and generated poor translation output. The reliance on rigid, handcrafted rules made it difficult for these systems to handle the nuances of language efficiently. Moreover, the lack of context-aware translation often led to awkward and inaccurate translations. Despite these shortcomings, the Georgetown IBM Experiment (1954) is one of the earliest attempts at machine translation, where researchers at Georgetown University and IBM developed an RBMT (Rule Based Machine Translation) system to translate Russian sentences into English using handcrafted linguistic rules as shown in Figure 3. This experiment marked a significant milestone, demonstrating that machine translation was possible which once was considered a human-only task, though still far from perfect. Another such example was Yamato Machine developed in Japan as shown in Figure 4, which focused on translating English to Japanese and served as an early demonstration of machine assisted translation for Asian languages. These pioneering efforts set the stage for further advancements, despite their initial limitations in accuracy and efficiency.

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

Figure 3: *IBM Machine-Aided Language Translation,* showcasing early computational efforts in translation technology [4]

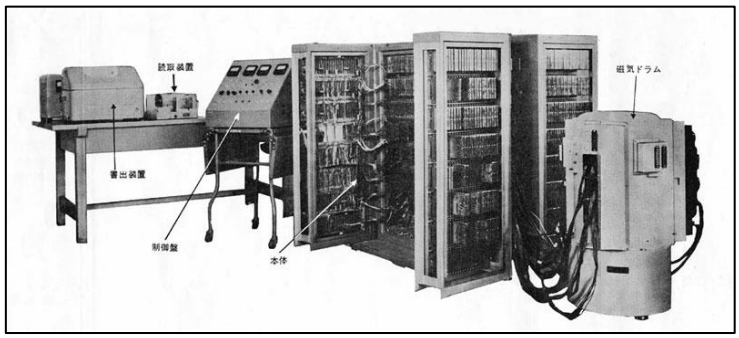

Figure 4: *Yamato Machine Translation System,* illustrating Japan's pioneering efforts in early machine translation technology [5]

2.4 SMT - Statistical Machine Translation

SMT played a pivotal role in the early machine translation period, offering a data-driven alternative to rule based approaches. These systems used statistical models to analyse large bilingual corpora and learn probabilistic mappings between source and target language phrases. The rise of SMT marked a significant shift from the manual, rule based translation methods to more automated, data-driven processes. These models typically included components for word alignment, phrase extraction, language modelling, and decoding. The translation process was heavily reliant on large parallel corpora, which served as the training data for the system. Yet they faced challenges, including data sparsity, domain adaptation issues, and the need for vast amounts of high quality bilingual corpora. Additionally, SMT struggled with handling long-range dependencies within sentences, which made it difficult to generate coherent translations. Capturing context-dependent meanings, and generating fluent, natural-sounding translations were also significant hurdles. Despite these challenges, SMT systems marked a significant advancement over earlier approaches as they reduced the reliance on handcrafted linguistic rules. One of the early successful commercial implementations of SMT was Candide system of IBM (1988) which helped establish the practicality of SMT for real-world applications. The Candide system was a milestone in machine translation, demonstrating how probabilistic models could be used to translate text efficiently, even with limited linguistic resources.

2.5 NMT - Neural Machine Translation

NMT utilized artificial neural networks to acquire the ability to directly translate from source languages to target languages. It revolutionized translation technology by not relying on predefined rules or alignments. It showed superior quality of result in comparison to the previous approaches, capturing syntactic nuances more accurately and producing fluent and natural sounding translations. NMT models, particularly those based on deep learning, are designed to learn from vast amounts of data, enabling them to handle a wide variety of linguistic patterns and structures. Nevertheless, these models necessitate substantial quantities of top-notch parallel data during training, which might be limited or absent for specific language pairs or domains, potentially resulting in biased data and leading to inaccuracies or mistranslations. These neural networks work by processing entire sentences at once, rather than word by word, allowing them to understand the broader context of a given sentence. It has also facilitated the creation of multilingual chatbots, virtual assistants, and learning tools. As a result, NMT systems can produce more coherent translations that maintain the integrity of the source text's meaning, even in complex scenarios. A research investigation [6] additionally revealed that translations generated by NMT, varying by book, were deemed to be on par in quality by native speakers of the target language compared to translations crafted by a

professional human translator, a result not observed with PBSMT, where the perception of quality varied at 8% and 20% respectively. Unlike earlier methods, which struggled with ambiguous terms and sentence structure, NMT can often choose the most contextually appropriate translation, leading to more accurate results. However, the challenges remain in areas such as low-resource languages, where the lack of sufficient data can hinder the system's ability to generalise and perform well. NMT has, yet, propelled machine translation closer to human-level performance and is widely used in real time translation applications such as social media platforms and global communication tools.

2.6 Transformer Models and Self-Attention Mechanisms

The Transformer architecture is a more advanced version of Neural Machine Translation (NMT), introduced in [7]. This architecture improves sequence-to-sequence learning, especially in machine translation, by using self-attention mechanisms to understand long-range dependencies and context. Unlike earlier models, the Transformer is highly parallelizable. This allows it to process much larger datasets efficiently and significantly reduces training time. The model relies on layers of attention mechanisms, where each word in a sentence is given different attention weights depending on its importance in context. This ability to capture complex relationships between words enables Transformers to generate translations that are both accurate and contextually rich. However, further work is needed to improve the model's interpretability, address biases, and ensure fairness and ethical considerations. Researchers are constantly working to mitigate the risk of harmful outputs such as reinforcing stereotypes or providing inaccurate translations in sensitive contexts. Additionally, ensuring that the model performs well across various languages particularly low-resource ones, remains an ongoing challenge. Despite these challenges, the Transformer architecture has rapidly become the standard for many state-of-the-art machine translation systems. It has opened the door to more sophisticated, multilingual applications, and continues to drive advancements in natural language processing.

3 Approach to Translations: Human, and Machine

Among the various technologies developed for translating literary texts, two approaches have primarily been highlighted: Human, and Machine. This reflects the ongoing tension between the expertise, creativity, and contextual understanding of human translators and the automation, efficiency, and scalability of machine translation systems. While human translators excel at producing nuanced, culturally sensitive translations, machine translation systems provide speed, accessibility, and continuous improvement complementing and enhancing human translation efforts. Machine translation systems, however, struggle with conveying idiomatic expressions and cultural references accurately, often requiring human intervention to ensure proper context is maintained. To achieve this most effectively, a combination of human expertise and machine assistance is necessary, leading to the need for a hybrid mode. The hybrid approach leverages the strengths of both methods to achieve high-quality translations that meet the diverse needs of the target audience.

4 Hybrid Mode in Translation

In context of literature, hybrid mode refers to the combined effort of both human translators and machine translation technology to produce translated or adapted literary texts. An initial translation is generated by the machine technology and is then edited and refined by human translators to retain the accuracy, fluency, and cultural appropriateness. Human translation experts enhance and refine the machine generated translation with the help of their linguistic expertise, cultural knowledge, and creative skills, to capture for the nuances, styles and context of the original literary work as explained in [8]. The hybrid approach enables

Proceedings DOI: 10.21467/proceedings.178
ISBN: 978-81-984081-8-1

Series: AIJR Proceedings
ISSN: 2582-3922

human translators to leverage the speed and scalability of machine translation while maintaining the depth of meaning and artistry that human translators bring. However, the approach still requires careful coordination to ensure that the machine-generated content aligns with the human translator's stylistic decisions. The advantages and challenges associated with this approach are considered as shown in the following section, where both the efficiency and the potential limitations are outlined.

4.1 Advantages of Hybrid Mode

Hybrid translation provides an advantage over other evolved technologies by improving productivity, cost effectiveness, and consistency. It delivers high quality translations with shorter turnaround times and reduced cost. It also reduces time and effort while maintaining the quality of translation. Human translation experts can hence focus on higher level tasks such as linguistic refinement, cultural adaptation, and creative expression preserving the original cultural essence of the text. While machine translation systems can handle routine translation tasks, provide and assist in generating initial drafts reducing time and increasing efficiency. This approach empathises the importance of ethical and cultural considerations in translating literature. Human translators play a crucial role in ensuring cultural heritage, author's voice and artistic vision in the texts drafted by the machine. This helps in correcting errors, clarifying ambiguities and corroborate respect of the author's work in the adaptation.

4.2 Challenges of Hybrid Mode

However, this technology still faces some challenges and is open to development. Managing the interaction between human translators and machine systems requires specialized and complex infrastructure. Ensuring consistency in the quality of work can be shown as a little challenging as human post editors can introduce a variation in style, tone or terminology. Allocating resources effectively, providing training and support, and fostering collaboration between human and machine translators are critical for the success of hybrid translation initiatives. Also. It may not always be cost effective, especially for large scale translations with limited resources. These models may struggle due to lack of domain knowledge and context required for correct and precise translation. Adaptation of machine outputs to specific domains and languages while ensuring accuracy and fluency is yet a challenge to be overcome.

5 Conclusion

The evolution of translation technology in literature has transitioned from manual efforts to highly sophisticated machine based systems. The hybrid mode offers to leverage human translation experts' strength in maintaining the artistic integrity and respect for author's voice and cultural heritage with machine's strength of automation and generating an initial draft saving time efficiently. Nevertheless, this technique also has many areas of development as mentioned above. A holistic approach is needed to address these challenges that combines technological innovation, organisational support, and human expertise. It also raises ethical and cultural considerations regarding authorship, intellectual property and cultural representation to respect the author's voice, cultural heritage and artistic integrity. It requires transparency, sensitivity and ethical considerate guidelines. By overcoming these stepping stones, this approach can pave the path for new doors for cross cultural communication, creative integrity and collaboration in global literary landscape.

6 Declarations

6.1 Competing Interests

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6.2 Publisher's Note

AIJR remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

How to Cite

Taruna Sharma, Nandini Sharma, Arpit Dwivedi, Daksh Thakur (2025). A Review of Major Translation Technologies in Literature: Bridging to Hybrid Systems. *AIJR Proceedings*, 64-70. https://doi.org/10.21467/proceedings.178.8

References

[1] A. Lange, C. Rundle, and D. Monticelli (eds.), "The Routledge Handbook of the History of Translation Studies": United Kingdom, Taylor & Francis, 2024.

[2] Encyclopaedia Britannica, "Biblical translation", *britannica.com*, photograph, 20 Jan. 2017. [Online]. Available: https://www.britannica.com/topic/biblical-translation. [Accessed Apr.15, 2024].

[3] G. Jansoone, "Dante Alighieri by Signorelli", *worldhistory.org*, 20 Mar. 2019. [Online].

Available: https://www.worldhistory.org/image/10260/dante-alighieri-by-signorelli/. [Accessed Apr. 27, 2024].

[4] IBM, "Machine-aided translation," photograph from "IBM Archive: Language Translation," [Video], *ibm*.com, n.d. [Online]. Available: https://www.ibm.com/history/machine-aided-translation. [Accessed Apr.15, 2024].

[5] IPSJ Computer Museum. "Dawn of Japanese Computers: FACOM 230-75 (1964).", *museum.ipsj.or.jp*, n.d. [Online]. Available: https://museum.ipsj.or.jp/en/computer/dawn/0027.html. [Accessed Apr. 15, 2024].

[6] A. Toral and A. Way, "What Level of Quality Can Neural Machine Translation Attain on Literary Text?" in *Translation Quality Assessment*: Cham, Switzerland, Springer, 2018.

[7] A. Vaswani, "Attention is All You Need," in *Advances in Neural Information Processing Systems*, in *31st Conference on Neural Information Processing Systems*, Long Beach, CA, USA, 2017, pp. 6000-6010. [Online] Available: https://user.phil.hhu.de/~cwurm/wp-content/uploads/2020/01/7181-attention-is-all-you-need.pdf

[8] Digital.gov, "Introduction to Translation Technology", *digital.gov*, n.d. [Online]. Available: https://digital.gov/resources/introduction-to-translation-technology/. [Accessed Apr. 18, 2024].