



The Ethics of Machine Translation in The Era of Artificial Intelligence

Wen Peng*

Wenhua College, Wuhan, Hubei, 430000, China

*corresponding author: 940793224@qq.com

Abstract: Along with technological progress, artificial intelligence has become a driving force for contemporary as well as future development. Artificial intelligence technology and humans are becoming more closely connected as it advances, and its influence on morals and ethics is also becoming more apparent. This paper examines explores the ethical dilemmas raised by machine translation technologies in the era of artificial intelligence. Initially, it presents an overview of the evolution of machine translation, with a focus on how neural machine translation has enhanced accuracy and efficiency. Subsequently, it delves into its applications across global communication, education, and other fields, alongside technological challenges such as data processing and model optimization. The core of the discussion on the ethical issues raised by machine translation, including data privacy, copyright disputes, accountability and cultural biases. In response to these challenges, the paper proposes solutions like improving data handling methods and enacting legal standards. The paper emphasizes that technological innovation and interdisciplinary collaboration are the only ways to establish a balanced approach to technological growth and ethical responsibility, and that both approaches are essential for the sound development of machine translation technology. This method has significant ramifications for the advancement of linguistic technologies in the era of globalization in addition to offering guidelines for the moral application of machine translation technologies.

Keywords: Machine Translation, Artificial Intelligence, Ethics

1 Introduction

The Weaver Memorandum, written by Warren Weaver and published in the 1950s, officially introduced the concept of machine translation [1]. After that, technological advancements facilitated the shift from rule-based translation methods to statistical machine translation, which was followed by the widespread use of neural machine translation based on deep learning [2]. The advent of neural machine translation marks a significant turning point in the field, resulting in substantial improvements in translation quality, particularly in terms of fluency and contextual consistency [3]. As artificial intelligence technologies continue to advance, machine trans-

lation is becoming increasingly important for international trade, cross-cultural communication, and information access [4].

As the rapidly development and widespread application of machine translation, its ethical issues have aroused people's concern, including data privacy, intellectual property rights, unemployment, algorithmic bias, and cultural invasion [5]. In the context of globalization, machine translation serves as a bridge between different languages and cultures [6]. Therefore, the solution of its ethical issues is of vital importance to a healthier development of machine translation, greater cultural diversity, and fairer individual rights.

2 Overview of Machine Translation Technology

Machine translation technology, as a core technology at the intersection of computer science and linguistics, aims to automatically translate a natural language text into another language through a computer program [7]. Machine translation technology has seen rapid development and significant transformation, moving from the early rule-based method to the later statistical machine translation to the now popular neural network-based approach. The flowchart of machine translation is shown in Figure 1. We shall discuss this procedure in the parts that follow.

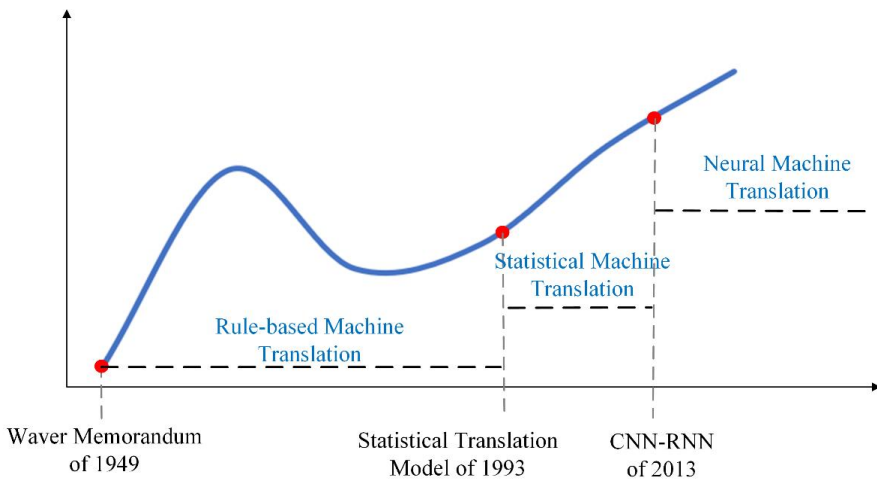


Fig. 1. Development progress of machine translation

2.1 Rule-based Machine Translation

One of the earliest machine translation methods, rule-based machine translation methods relied heavily on rules and dictionaries manually written by linguists to translate words and phrases from the source language into the target language. Rule-based machine translation (RBMT) relies on complex grammatical and lexical rules established by linguists [8]. This method can translate text in a grammatically correct manner. however, it is extremely time-

consuming and inefficient, and difficult to cover all the nuances of the language due to the limitation of the rules.

2.2 Statistical Machine Translation

Statistical machine translation (SMT) relies on a large bilingual corpus to learn translation rules, and the core idea is to translate by analyzing statistical relationships between source and target languages [9]. The advantage of this approach is that it can deal with complex linguistic text that are challenging for rule-based approaches to handle. However, its main limitations are its reliance on large-scale parallel corpora, poor fluency and readability.

2.3 Neural Machine Translation

A significant advancement in machine translation technology is neural machine translation (NMT), which translates text end-to-end by utilizing deep learning techniques, particularly Sequence to Sequence models and attention mechanisms [10]. The advantage of this approach lies in its ability to produce texts that are more fluid and natural. Thanks to its unique neural network architecture, translation quality and efficiency are greatly enhanced.

Among the ethical issues of machine translation, one of the most significant concerns is the problem of mistranslation. To identify the causes of this problem, there are three steps in the process of using large-scale language models. The first step involves the preprocessing of original data, including labeling and data gathering. In the context of deep learning, the second step primarily focuses on feature engineering and data training. The third phase integrates the trained big language model into application software [11]. Our research has shown that the primary causes of mistranslations occur during the preprocessing stage of the data, especially during feature engineering when there is inaccurate labeling. To minimize the possibility of algorithmic bias in machine translation, it is crucial to develop a data discrimination and filtering model that removes information containing bias related to politics, gender, and race.

Over the recent years, with the rapid development in the field of artificial intelligence, especially deep learning technology, the quality of machine translation has been significantly improved. In a word, the development of machine translation not only demonstrates the advances in computer science, but also reflects the potential of language processing technologies in understanding and generating natural language. By gaining a deeper understanding of these technical principles and their development, we can better grasp the future trends and challenges of machine translation.

3 Ethical Issues in Machine Translation

With the widespread use of machine translation technology, ethical problems are becoming more pressing and need to be addressed. Several key ethical issues are discussed below:

Data Privacy and Security: Machine translation systems often have to process texts that contain sensitive or private data. The primary issue that machine translation technology needs to solve is how to ensure that these data are neither misused or disclosed throughout the translation process. To protect users' privacy and security, a number of steps should be taken, such as improving data encryption technology, restricting access to sensitive information, and anonymisation.

Copyright Issue: The copyright issues involved in machine translation mainly revolve around the copyright ownership of the translated work [12]. Especially when translating literary works, professional articles and other contents, translation without the permission of the original author may infringe the copyright of the original work. Therefore, the rules that machine translation must go by are defining who owns the copyright to translated works and protecting the original authors' intellectual property rights.

Responsibility and reliability: Errors in machine translation can have detrimental effects, particularly when they happen in critical domains like legal and medical documents. At this point, defining responsibility becomes a complex issue. In order to ensure that issues can be handled promptly and efficiently when they happen, machine translation service providers must not only improve the accuracy and reliability of their translations but also set up a clear responsibility attribution mechanism.

Culture and prejudice: Machine translation may exacerbate cultural biases and inequities when processing texts from different cultural contexts. This is because machine translation models are often trained based on existing datasets that may contain specific cultural biases. Therefore, eliminating data bias and ensuring cultural diversity in the translation process are important tasks in the development of machine translation technology.

To sum up, with the continuous progress of machine translation technology and the expansion of its application fields, its ethical issues need urgent attention and solution. Through interdisciplinary cooperation and the formulation of reasonable policies and guidelines, these challenges can be effectively addressed to promote the sound development of machine translation technology.

4 Ethical Principles and Solutions

In the field of machine translation, the resolution of ethical issues requires a combination of technical, legal and multidisciplinary cooperation. To address issues such as data privacy, copyright, definition of liability, cultural bias and algorithmic transparency, this chapter proposes a series of solutions and explores ways to verify the effectiveness of these solutions.

Technical Solutions: To address the ethical issues, the architecture, algorithms and data processing methods of machine translation models can be improved. For example, data encryption methods are used to protect user data—especially sensitive data—from being misused. Through copyright information tagging and copyright sharing agreements, algorithms can be designed to automatically recognize and respect the original work's copyright while handling copyright issues. To address the issue of cultural bias, datasets should be optimized.

Laws and Social Policies: The development of relevant legal norms and social policies are critical to addressing the ethical challenges raised by machine translation. This includes the creation of data protection laws to ensure the security of personal information; the implementation copyright law reform to adapt to the ways of creating and sharing information; and the development of AI ethical standards and guidelines to clarify the ethical boundaries of technology development and application.

Multidisciplinary cooperation: Policymakers, linguists, ethicists, and scientists must work closely together to address the ethical concerns raised by machine translation. Such interdisciplinary collaboration can help to understand and address the ethical challenges of technology in a holistic manner, and jointly contribute to the implementation of effective measures.

5 Conclusion

This study systematically explores the ethical challenges raised by machine translation technology in the age of artificial intelligence, including data privacy and security, copyright issues, liability and reliability, and culture and bias. Through in-depth analyses, we find that although machine translation technology provides unprecedented convenience for cross-linguistic communication, it also raises a series of ethical issues. If we left these issues unaddressed, it may impede the sound development of the technology and harm user trust. This study proposes a set of solutions to addressing these ethical challenges through technological improvements, the development of relevant legal norms and social policies, and interdisciplinary cooperation. The study believe that machine translation will gain great progress and better serve people in the future.

Reference

1. Melby, A. K. (2019). Future of machine translation: Musings on Weaver's memo. *The Routledge handbook of translation and technology*, 419-436.
2. Ali, M. N. Y., Rahman, M. L., Chaki, J., Dey, N., & Santosh, K. C. (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), 2365-2376.
3. Freitag, M., Foster, G., Grangier, D., Ratnakar, V., Tan, Q., & Macherey, W. (2021). Experts, errors, and context: A large-scale study of human evaluation for machine translation. *Transactions of the Association for Computational Linguistics*, 9, 1460-1474.
4. Fortunati, L., Katz, J. E., & Riccini, R. (Eds.). (2003). *Mediating the human body: Technology, communication, and fashion*. Routledge.
5. Taivalkoski-Shilov, K. (2019). Ethical issues regarding machine (-assisted) translation of literary texts. *Perspectives*, 27(5), 689-703.
6. Lopez, A. (2008). Statistical machine translation. *ACM Computing Surveys (CSUR)*, 40(3), 1-49.
7. Xie, S. (2018). Translation and globalization. In *The Routledge handbook of translation and politics* (pp. 79-94). Routledge.

8. Shiwen, Y., & Xiaojing, B. (2014). Rule-based machine translation. In Routledge encyclopedia of translation technology (pp. 186-200). Routledge.
9. Koehn, P. (2009). Statistical machine translation. Cambridge University Press.
10. Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
11. Mikolov, T., Deoras, A., Povey, D., Burget, L., & Černocký, J. (2011, December). Strategies for training large scale neural network language models. In 2011 IEEE Workshop on Automatic Speech Recognition & Understanding (pp. 196-201). IEEE.
12. Moorkens, J., & Lewis, D. (2019). Copyright and the re-use of translation as data. In The Routledge handbook of translation and technology (pp. 469-481). Routledge.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

