Journal of Artificial Intelligence Practice (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/jaip.2023.060706
ISSN 2371-8412 Vol. 6 Num. 7

Exploration of Computer-Assisted Translation Technology in Translating Technical Terms in Traditional Chinese Medicine under the Perspective of AI Vision

Yang Zheng[1], Su Yifu[2], Tan Xiumin[1,*]

[1]*Tianjin University of Traditional Chinese Medicine, Tianjin, China*
[2]*The Second Nankai High School in Tianjin, Tianjin, China*
[*]*Corresponding author: tanxiumin@163.com*

Keywords: AI, Computer Aided Translation, Academic language in Traditional Chinese Medicine

Abstract: With the recent surge of AI technology, Computer Aided Translation (CAT) has played as a backbone in TCM translation. This paper introduces CAT technology from the perspective of AI and features of academic language in Traditional Chinese Medicine, and analyzes the advantages and problems of AI+CAT in Chinese medicine translation and related corpus construction. Through the study of this paper, we can better understand the application prospect of computer-aided technology in Chinese medicine translation in the era of AI and provide useful reference for further research in related fields.

1. Introduction

Computer assisted translation (CAT) is of great application value in the academic translation of Chinese medicine. As a treasure of Chinese culture, TCM has attracted much attention from the global medical community for its unique theoretical system and diagnosis and treatment methods. However, because Chinese medicine involves a large number of unique concepts and terms, its translation has a high degree of difficulty and challenge. In the era of artificial intelligence, how to use computer-aided translation technology to solve this problem and promote the international dissemination of traditional Chinese medicine culture is a hot issue in current research. This paper first expounds the concept of computer-assisted translation technology and the characteristics of TCM academic language from the perspective of AI, then discusses the advantages of AI+CAT in TCM academic language translation, and finally analyzes the existing problems in the construction of TCM terminology corpus under the condition of AI computer-aided integration, hoping to provide references for relevant researchers.

2. Computer-Assisted Translation Technology in the AI Perspective

Computer-assisted Translation Technology (CAT) in the AI perspective refers to a technology that assists human translation through artificial intelligence technology and computer algorithms. On the

one hand, AI-assisted translation tools can combine machine translation technology to provide translation suggestions and candidate translation results that translation professionals need, thus improving work efficiency by establishing language models and probability models through AI algorithms. On the other hand, computer-assisted translation technology in the AI perspective can provide more accurate translation results through natural language processing and semantic analysis techniques. By using AI deep learning algorithms, the meaning and expression of the translated text can be better understood by analyzing the semantics and context of the translated text. Assisted translation tools can provide accurate translation suggestions based on context and semantic understanding and continuously learn and optimize translation results through feedback mechanisms, thus improving translation quality [1].

3. The Characteristics of the Academic Language of Traditional Chinese Medicine

The characteristics of the academic language of Traditional Chinese Medicine (TCM) are mainly reflected in the richness and diversity of terminology. Firstly, the abundance and diversity of terms in TCM reflect its profound history and extensive knowledge system. TCM originates from ancient ethnic cultures and has developed over thousands of years, forming a vast and complex theoretical system. It has created a large number of terms and specialized vocabulary to express different conditions, pathologies, treatment methods, medications, and more. These terms and specialized vocabulary not only have rich connotations but also accurately convey specific medical information. Secondly, the diversity and abundance of terms in reflect its personalized observation and treatment of disease phenomena. Unlike Western Medicine, which focuses on the study of pathology and physiology, TCM pays more attention to comprehensive analysis and judgment of disease symptoms, constitutional characteristics, and etiology differentiation. For example, "qi deficiency" refers to insufficient qi and blood, indicating physical weakness; "blood stasis" refers to poor blood circulation, leading to blood accumulation and causing pain and other symptoms; "phlegm-dampness" refers to the mixture of dampness and phlegm, which can easily lead to respiratory difficulties, coughing, chest tightness, and other symptoms. These terms reflect the observation methods and disease classification system of traditional Chinese medicine, using personalized observation and analysis of disease phenomena and patient symptoms for diagnosis and treatment[2].

The characteristic of the academic language in TCM is also manifested in vivid metaphorical and symbolic language. Metaphors and symbols are often used in the academic terminology of TCM to describe disease phenomena, etiology and pathogenesis, and treatment principles. This is determined by the characteristics of TCM and its cultural background. For example, "insufficient qi and blood" is metaphorically described as "a river without flow, a stream without water"; "invasion of cold evil" is metaphorically described as "a cold north wind blowing on flower petals"; "disharmony between the heart and spleen" is metaphorically described as "wood and fire unable to support each other," and so on[3]. These metaphors and symbolic language enrich the imagery and emotional color of the academic terminology of traditional Chinese medicine, making it more inspiring and memorable.

TCM academic language is delicate and specific. TCM academic language often provides detailed and specific descriptions of diseases, etiology and pathogenesis, diagnostic methods, and treatment principles, which is one of the characteristics of TCM. In terms of disease description, TCM academic language usually includes detailed descriptions of the symptoms, manifestations, and development process of diseases. In terms of etiology and pathogenesis, TCM academic language often provides detailed explanations of the causes and pathological mechanisms of diseases. TCM believes that the occurrence of diseases is related to the imbalance of the internal and external environments of the human body, so the description of etiology and pathogenesis generally includes internal and external factors, as well as the physiological and pathological mechanisms of the disease [4]. For example, in

TCM, the etiology and pathogenesis of a common cold may be explained as the invasion of external wind cold or wind heat evil into the body, which leads to the imbalance of the body's defense function and thus triggers a cold. Furthermore, TCM also provides detailed descriptions of diagnostic methods. TCM academic language contains rich diagnostic terms, including names and descriptions of various diseases, such as "qi stagnation," "damp heat," "phlegm dampness," etc. These terms can accurately express the characteristics and changes of the disease, which helps doctors make diagnoses. TCM academic language also provides detailed descriptions of diagnostic methods. The main diagnostic methods in TCM are the Four Diagnostic Methods of observation, listening and smelling, questioning, and palpation. These methods are described in detail in TCM academic language, including observing the external characteristics of the patient, listening to their voice and smell, inquiring about their symptoms and medical history, and performing palpation on the pulse and tongue coating, in order to conduct comprehensive analysis and diagnosis.

4. The Advantages of AI+CAT in the Translation of Traditional Chinese Medicine Terminology

CAT tools are technological tools that help translation professionals improve translation efficiency and quality. AI-assisted CAT tools can ensure improved machine and human translation efficiency in the progressive development of artificial intelligence, with the following implications:

4.1. Terminology management

CAT tools typically have terminology management features, allowing for the collection, storage, and management of professional terminology used in translation projects. In AI-assisted machine translation processes, artificial intelligence can match potential terms in computer texts with CAT terminology databases through algorithms and learning, ensuring consistent translation guidelines and uniform expressions.

4.2. Terminology consistency checking

AI, through algorithms and learning, can automatically identify and mark translated terms in texts, while providing consistency checking functions to ensure the use of the same terms throughout the project. This helps avoid randomization and non-standardization of existing CAT terms in the same context, maintaining accuracy and uniformity in professional terminology.

4.3. Translation Memory

CAT translation tools are based on the functionality of artificial intelligence memory banks, which can record and learn from translators' previous translation content and provide matching and revision suggestions during subsequent translation processes. When similar or repetitive texts appear, CAT tools assisted by artificial intelligence can automatically extract previously translated memory paragraphs, reducing workload and translation redundancy, while helping to maintain consistency in terminology translation.

4.4. Terminology Reports and Statistical Analysis

CAT provides a range of functions during the translation process, including Traditional Chinese Medicine (TCM) terminology reporting and statistical analysis. CAT tools usually provide terminology database functions, creating and managing emerging terminology data with the support of AI machine learning capabilities. Artificial intelligence interactive software such as Google,

DeepL, ChatGPT, etc., can help translators automatically identify and extract TCM terminology during the translation process, and summarize them to generate terminology reports. This allows translators to easily search for and check the usage of TCM terminology, improving translation quality and consistency. Moreover, CAT tools can provide statistical analysis functionality for TCM terminology by recording the terms and translation results used by translators during the translation process, generating statistical data. By overlaying CAT statistical analysis with AI interactive software, translators can understand the frequency of TCM terminology usage, common translation issues, terminology consistency, and other information. These statistical data can help translators better grasp the usage standards and characteristics of TCM terminology, improving translation efficiency and accuracy[5].

4.5. Parallel text alignment and comparison

CAT tools can automatically align the source text and the translated text, displaying them simultaneously in one interface. This allows translators to visually see the correspondence between the source text and the translation, making it easier to identify potential errors and differences. AI-powered interactive software can automatically detect repeated text in the translation and provide relevant suggestions and paragraphs using the previously translated content from the translation memory. This eliminates the need for translators to translate the same or similar content repeatedly, greatly reducing their workload. Moreover, the use of translation memory ensures consistency in the translation, especially in the translation of long documents or multiple documents.

4.6. Automatic translation and machine translation

CAT tools can integrate AI-powered translation engines to automatically translate text. Although the accuracy of machine translation may be relatively limited in the field of traditional Chinese medicine, it still has its value and role in AI-assisted machine translation processes. Machine translation engines may provide accurate translation results for common terms, phrases, or fixed expressions, serving as a reference for translators[6]. This helps translators quickly determine translation choices in certain contexts, improving translation efficiency. Additionally, when using CAT tools for translation, translators can compare the source text with the machine-translated results, quickly identifying differences and issues and making necessary modifications. This avoids starting translation from scratch, saving a significant amount of time and effort. With the continuous advancement of machine translation technology, especially in the development of AI-based neural network machine translation, machine translation may play an even greater role in the translation of traditional Chinese medicine in the future. Neural network machine translation, through extensive training data and deep learning algorithms, can simulate the human brain's translation process, enhancing translation quality and accuracy. With the improvement of neural network machine translation, it can be foreseen that machine translation results may become more accurate and reliable in the field of traditional Chinese medicine, particularly for common terms and sentence patterns.

5. The Problems in the Construction of TCM Terminology Corpus under the Integration of AI and Computer-Aided Technology

5.1. The issue of standardization and accuracy of terminology

The standardization of terminology refers to the use of consistent terms for the same concept or phenomenon in the field of traditional Chinese medicine, ensuring smooth communication and dialogue between different schools and regions. The accuracy of terminology means that the

definitions and use of terms should be precise and free of ambiguity, ensuring the rigor and reliability of academic research. However, the development and evolution of traditional Chinese medicine have a long history, resulting in differences in the understanding and application of terminology among different regions and schools. These differences may lead to difficulties in academic communication, hence the need to establish a common terminology database to address this issue. To address these problems, it is necessary to establish an integrated traditional Chinese medicine terminology database that incorporates AI and computer-assisted technologies. Firstly, professionals need to conduct research and analysis on the terminology used in different regions and schools with the help of AI algorithms and statistical tools, collecting the characteristics and definitions of terminologies from all parties to ensure comprehensive coverage. Secondly, through expert discussions and academic conferences, multiple opinions can be incorporated to reach a consensus on the definition and usage norms of terminology. In addition, through continuous standardization adjustments of traditional Chinese medicine terminologies in the database using AI and computer-assisted technologies, standardized settings can be established for the definition and usage of terminologies. This will help enhance the accuracy and standardization of traditional Chinese medicine terminology, avoiding misunderstandings and confusion[7].

5.2. Updating and improving terminology

Traditional Chinese medicine is constantly developing and evolving, with new theories and treatment methods emerging. Therefore, both the AI's Chinese medical terminology interaction function and the CAT's Chinese medicine terminology bank function need to be continuously updated and improved to adapt to new developments. This requires ongoing research and monitoring on the basis of establishing a terminology bank, promptly incorporating new terms into the bank.

As new terms are continuously introduced, it is also important to improve the terminology bank. The improvement of the terminology bank includes the machine learning function of artificial intelligence and the neural network machine translation function. It involves real-time monitoring and inclusion of more accurate and precise terms and their definitions. This helps improve the clarity and standardization of Chinese medicine terminology, avoid confusion and misunderstandings, and promote the development of the discipline[8].

To achieve the updating and improvement of the terminology bank, it is necessary to establish a continuous research mechanism and monitoring mechanism. This can be done through collaboration with Chinese medicine research institutions, clinical doctors, AI and computer-aided translation research institutions, etc., to establish specialized terminology research teams. These teams should regularly study and sort out new knowledge and terminology in the field of Chinese medicine to ensure the updating and improvement of the terminology bank. Additionally, establishing a dedicated review mechanism to review and confirm new terms and definitions is essential to ensure the accuracy and standardization of the terminology.

5.3. Accessibility and usability

The establishment of the translation function of AI integrated with CAT in traditional Chinese medicine terminology requires its integration into existing academic resource systems and providing interfaces that are easy to access and use. Firstly, the accessibility of the traditional Chinese medicine terminology database refers to whether the academic terminology database is easily accessible and browsable. With the development of artificial intelligence and the internet, establishing an online traditional Chinese medicine terminology database can facilitate researchers, clinical doctors, and students to query the definitions and explanations of relevant terms anytime and anywhere, thereby improving learning and research efficiency. However, it is currently difficult to find an online AI

integrated traditional Chinese medicine terminology database in China. Secondly, the usability of the traditional Chinese medicine terminology database refers to whether the academic terminology database is easy to use and understand. For example, artificial intelligence interaction tools require accurate instructions and steps as guidance in order to generate relevant answers and translations related to the topic. This places high demands on users' proficiency in the source language and target language and their accurate usage. Moreover, traditional Chinese medicine terminology has its own unique characteristics and connotations, often requiring a certain level of professional knowledge to understand and apply. Therefore, a good academic terminology database should provide detailed explanations and examples to help users better understand and apply these terms. In addition, the academic terminology database should also have a well-organized structure and search function, making it convenient for users to search for and use related terms according to different needs.

6. Conclusion

Translation of traditional Chinese medicine (TCM) is a tedious and challenging task, and traditional translation methods are no longer able to meet the requirements of modern society for high-quality translation results. Computer-assisted technologies in the AI field provide a new solution for TCM translation. By utilizing AI's machine learning, neural network translation, and natural language processing technologies, large amounts of TCM literature and knowledge can be organized and analyzed to generate high-quality machine-assisted translation results with improved accuracy and consistency. This not only greatly accelerates the speed of TCM translation, but also enhances the quality and precision of translation. However, computer-assisted technologies in the AI field still face some challenges and issues in TCM translation. The complexity of TCM translation arises from the uniqueness of TCM, such as the need to combine TCM theories and practices in translating medicinal names and treatment methods. Differences in language and cultural backgrounds also need to be taken into consideration, as TCM is primarily expressed in Chinese. Additionally, the continuous development of AI technology requires us to continuously improve and optimize algorithms and models in TCM translation to adapt to new translation needs.

References

[1] Duan Yingshuai, Duan Yishan. Research status and analysis of TCM terminology translation in recent 30 years [J]. Chinese Journal of Traditional Chinese Medicine, 2019, 35(01):442-446.

[2] He Yanqiu, Tu Yuming. English Translation of Traditional Chinese Medicine from the perspective of cross-cultural Communication [J]. English Plaza, 2020(34):45-48.

[3] Yang, Y., Xiong, J., Wang, C., et al. (2018). Comparison between Artificial Intelligence Thinking Mode and Traditional Chinese Medicine "Image Thinking". Chinese Journal of Traditional Chinese Medicine, 33(10), 4420-4421.

[4] Li, Z., & Kong, L. (2020). Teaching Practice of Chinese Medicine Translation: Taking Cultural-Loaded Terms in "Shang Han Lun" as an Example. Science Education Tribune, 30(12), 181-182.

[5] Li, F. (2022). A Comparative Study on the Quality of Man-Machine English-Chinese Translation in the Era of Artificial Intelligence. Foreign Language World, 2022(04), 74-78.

[6] Du, M., & Lu, H. (2018). Bilingual Alignment Problems and Countermeasures in the Compilation of Translation Memory. Overseas English, 2018(06), 125-126.

[7] Shao, H., & Yang, X. (2018). Current Situation and Development of Artificial Intelligence language Translation. Electronic World, 2018(16), 28-29.

[8] Zhang, J. (2023). Application of Computer-Aided Translation in Learner-Centered English Translation Teaching of Traditional Chinese Medicine. Western Chinese Medicine, 36(1), 45-47.