

A Comparative Study on the Effectiveness of ChatGPT and DeepSeek in the English Translation of Characteristic Terms from Classical Chinese Medical Texts—A Case Study of Huangdi Neijing: Jue Lun Chapter 45

Li Yuhang^{1*}

¹ Xi'an Polytechnic University, Shaanxi, China

* Correspondence: m18091325699@163.com

<https://doi.org/10.53104/insights.soc.sci.2026.06006>

Received: 3 June 2026

Revised: 8 June 2026

Accepted: 12 June 2026

Published: 16 June 2026

Citation: Li, Y. (2026). A Comparative Study on the Effectiveness of ChatGPT and DeepSeek in the English Translation of Characteristic Terms from Classical Chinese Medical Texts—A Case Study of Huangdi Neijing: Jue Lun Chapter 45. *Insights in Social Science*, 4(2), 66-75.

Copyright: © 2026 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: This study employs Chapter 45 (Jue Lun) of Huangdi Neijing (Yellow Emperor's Inner Classic) as the source text to compare the performance of ChatGPT and DeepSeek in the English translation of characteristic terms from traditional Chinese medicine (TCM) classics. Results indicate that while ChatGPT produces fluent translations with strong cross-contextual adaptability, it exhibits terminological inconsistencies and a tendency to dilute deep cultural connotations. In contrast, DeepSeek, leveraging a domain-specific knowledge base, demonstrates greater stability in recognizing the underlying medical rationale of core terminologies, thereby generating translations with higher academic rigor. However, neither model fully reconciles the competing demands of terminological precision, textual readability, and philosophical depth. Consequently, this study proposes a "Generation-Proofreading-Optimization" human-AI collaborative translation model. This approach integrates the technological strengths of both systems while emphasizing the translator's pivotal role in strategy formulation and cultural gatekeeping, offering a viable pathway for the intelligent translation and international dissemination of TCM classics.

Keywords: ChatGPT; DeepSeek; translation of TCM classics; terminology translation; human-AI collaboration

1. Introduction

In recent years, the rapid development of large language models (LLMs) has driven continuous breakthroughs in artificial intelligence technology and its applications (Qin & Zhou, 2024). ChatGPT

and DeepSeek, both based on LLMs, possess distinct advantages in deep learning and natural language processing. They represent two paradigms of translation technology: general-purpose and domain-specific models, bringing new

opportunities and challenges to the translation industry. As a newly launched AI system in China, DeepSeek's translation capabilities have garnered attention from researchers and learners. Currently, domestic research on DeepSeek's translation performance remains limited, particularly regarding its effectiveness in translating classical texts of traditional Chinese medicine (TCM).

Based on ChatGPT-4 and DeepSeek V3, this study selects Chapter 45 (Jue Lun) of the Huangdi Neijing (Yellow Emperor's Inner Classic) as the corpus. It analyzes the linguistic differences between the two major language models to address three core questions: What is the quality of translations generated by these AI translation tools? Which tool can better avoid terminological mistranslations and deviations? And which tool is more capable of conveying philosophical connotations? This paper aims to provide empirical evidence for the feasibility of applying AI in translating characteristic terminology of classical TCM texts.

2. Research Design

2.1 Corpus Acquisition

The quality of prompts directly determines the output performance of LLMs. Well-crafted instructions can improve the accuracy and coherence of the results while optimizing model

performance (Chen et al., 2023; Sahoo et al., 2024). To enhance the targetedness and coherence of model translation, this study set restrictive instructions when inputting the source text of the Jue Lun chapter into the GPT-4 and DeepSeek V3 models: "Please translate the following professional text from the Huangdi Neijing into English, accurately conveying the cultural connotations of TCM terminology: + [Text]." Each model independently generated three translations, and the first output was recorded as the primary sample for analysis. This choice was made because the first output is closest to the model's direct response to the same prompt, before any additional human selection or adjustment is introduced. The other two outputs were used as supplementary references to observe output variation, but they were not treated as independent samples for statistical comparison. It should also be noted that LLM-generated translations may vary across repeated outputs, and the use of the first output as the main analytical sample is therefore a limitation of this study.

Terminology translation is a cross-cultural communicative activity, with the primary goal being the accurate transmission of conceptual meaning (Zhang et al., 2025). Initially, terminology extraction tools (such as Trados and Yufan Terminology) were employed to extract key terms from the source text, yielding the results shown in Table 1.

Table 1. Key Terminology Extraction Table

No.	Name	Frequency	Definition	Source Text Example
1	Jue	25	Pathological state caused by disordered Qi flow	"厥之寒热者，何也？"
2	Qi	13	Basic subtle substance constituting life	"气因于中，阳气衰，不能渗营其经络"
3	Xu	9	Pathological state of insufficiency or functional decline of healthy Qi, leading to weakened physiological function and disease resistance	"虚则补之，不盛不虚"
4	Yang Qi	8	Refined Qi with warming and promoting functions	"阳气衰，不能渗营其经络"
5	Yin Qi	8	Refined Qi with cooling and calming functions	"阴气起于足五指之里"
6	Re Jue	4	Syncope syndrome caused by Yin deficiency and Yang hyperactivity	"阴气衰于下，则为热厥"
7	Han Jue	4	Syncope syndrome caused by Yang deficiency and Yin excess	"阳气衰于下，则为寒厥"

8	Taiyin	4	Meridian name, referring to the Spleen or Lung Meridian	“太阴厥逆，腑急挛”
9	Yangming	4	Meridian name, referring to the Stomach or Large Intestine Meridian	“阳明厥逆，喘咳身热”
10	Jingluo	4	The channel system through which Qi and blood circulate	“不能渗营其经络”
11	Shouzu	3	Limbs/Hands and feet, common symptom locations for syncope syndrome	“使人手足寒，三日死”
12	Fuzhang	3	Abdominal distension due to Qi stagnation, fluid retention, or deficient Qi obstruction	“腹胀，泾漫不利”
13	Fuman	3	Abdominal fullness due to disordered and obstructed Qi flow in the abdomen	“腹满不得卧，面赤而热”

2.2 Analytical Framework

The primary criterion for terminology translation is accuracy. As “terminology is the core and soul of specialized language” (Jiang & Wu, 2019), the accuracy of TCM terminology translation directly relates to the scientific nature of TCM’s cross-cultural dissemination. Therefore, this study regards it as the most crucial evaluation metric. Furthermore, conciseness is another important principle. Li (2008) noted, “A defining characteristic of TCM terminology is its brevity, which should be maintained in English translation.” However, mere terminological translation cannot fully capture the ambiguous explanations and metaphorical expressions inherent in TCM (Zhu & Wang, 2025). Consequently, translators processing TCM texts must heavily rely on the context, adopting differentiated translation strategies to facilitate reader comprehension and improve the fluency and readability of the translation.

This study compares the translation performance of the extracted terms across three dimensions: accuracy, conciseness, and readability.

3. Results and Analysis

3.1 Accuracy

This study references the WHO International Standard Terminologies on Traditional Medicine in the Western Pacific Region (hereinafter referred to as Standard 1) to evaluate terminology translation accuracy. Given the expressional differences

between the extracted terms and those in Standard 1, translations aligning with the “English term” or “Synonyms” in the standard are deemed accurate during retrieval. The results are shown in Table 2. Zong (2020) pointed out that with sufficient training corpora, machine translation accuracy for domain-specific texts typically reaches over 75%. Examining the translations of the 13 core terms extracted in this study, both GPT-4 and DeepSeek V3 achieved a high baseline accuracy rate. Specifically, GPT-4 exhibited mistranslations deviating from the specific TCM context in three instances (Jingluo, Shouzu, Fuman); whereas DeepSeek V3 effectively circumvented the aforementioned anatomical generalization pitfalls, showing deviations in only two instances (Jingluo, Fuman). This micro-level discrepancy suggests that, in the selected examples, DeepSeek V3 shows relatively greater stability in adhering to TCM medical principles and aligning professional terms. However, this finding should be understood within the limited corpus of Chapter 45 rather than as a general conclusion about all TCM translation tasks. Notably, when translating the term “Shouzu” (手足), the GPT-4 model outputted “limbs.” This word is an anatomical term in Western medicine, defined as “appendages” in the Cambridge Dictionary. Since Standard 1 explicitly requires prioritizing literal translation of source language concepts and only resorting to free translation when literal translation is impossible, “Shouzu” is an unambiguous term suitable for literal translation. GPT-4’s rendering violates this principle. Moreover,

in TCM theory, “Shouzu” refers not merely to anatomical parts but serves as a key element in syndrome differentiation. Retaining the

independent semantics of “hands/feet” through literal translation avoids the ambiguity “limbs” brings to precise TCM diagnosis.

Table 2. Terminology Translation Consistency Comparison

LLM Model	Consistently Translated Terms	Inconsistently Translated Terms	Count
GPT-4	Jue, Qi, Xu, Yin Qi, Yang Qi, Han Jue, Re Jue, Taiyin, Yangming, Fuzhang	Jingluo, Shouzu, Fuman	About 77%
DeepSeek V3	Jue, Qi, Xu, Yin Qi, Yang Qi, Han Jue, Re Jue, Taiyin, Yangming, Fuzhang, Shouzu	Jingluo, Fuman	About 84.6%

3.2 Conciseness

TCM often utilizes concise and condensed language to express complex medical logic. Therefore, whether the translation retains this brevity is a crucial criterion for assessing quality.

Example 1 黄帝问曰：“厥之寒热者，何也？”

ChatGPT Translation: The Yellow Emperor asked: “What causes the manifestations of cold and heat in **Syncope (reversal syndromes)**?”

DeepSeek Translation: Huangdi asked: “What causes the cold and heat in **Syncope**?”

The above translation example demonstrates that ChatGPT utilized the core term “Syncope” while adding a supplementary explanation “(reversal syndromes)” in parentheses. Although this approach might intend to elucidate the meaning of Jue (厥), since the target audience primarily consists of individuals with medical backgrounds and the standard corresponding term for Jue in TCM is sufficiently clear, excessive elaboration conversely makes the translation verbose and increases the reading burden. In contrast, DeepSeek’s handling is concise and straightforward without extraneous annotations, adhering to the principle of conciseness in TCM terminology translation.

Example 2 阳气衰于下则为寒厥

ChatGPT Translation: When **yang qi (active energy)** declines below, cold syncope occurs due to imbalance.

DeepSeek Translation: When **Yang Qi** declines below, it results in Cold Syncope.

Since the core TCM term “Yang Qi” is an internationally accepted translation in academia,

DeepSeek directly translated it as “Yang Qi,” ensuring conciseness while complying with translation norms. Although ChatGPT also used “yang qi,” the supplementary “active energy” in parentheses is an unnecessary addition. The reason is that the concept of “Yang Qi” in TCM is complex, encompassing functions such as warming, promoting, and exciting; simply equating it to “active energy” is inaccurate and prone to ambiguity.

These instances indicate that AI tends to over-explain specialized terminology. For instance, when translating sentences containing DNA, it might explain it as “deoxyribonucleic acid” (Hesi Wang & Ma, 2023). However, for terms with standard equivalents that are common in medical texts, human translation typically uses the standard term directly without excessive explanation. In these examples, over-explanation may reflect a limitation in the models’ handling of standardized domain terminology. It can weaken the pragmatic independence of TCM terms and may also introduce unnecessary interpretive guidance for readers.

3.3 Readability of the Translation

As a high-context language, the meaning of words and characters in Chinese is deeply influenced by the surrounding context. Translating TCM texts requires analyzing the deep semantics of terms and sentences within specific contexts to avoid rigid word-for-word translation, while simultaneously ensuring fluency and readability.

Li Zhaoguo is a practitioner and theoretical researcher of the English translation of the Huangdi

Neijing in China. The International Standard Chinese-English Basic Nomenclature of Chinese Medicine (2007) by the National Administration of Traditional Chinese Medicine adopted 83% of his terminological translations. Thus, Li's translation wields significant influence, representing the interpretation of national classics by bilingual translators (Gong & He, 2022). Paul U. Unschuld (2011) is a renowned German medical historian and translator. Judging by the global library collections of various English editions of the Huangdi Neijing, Unschuld's version is the most widely collected (Wei Wang & Qu, 2024), making it a highly esteemed translation in Western countries. The difference lies in their approaches: Li tends to start from traditional Chinese culture (i.e., the pragmatic perspective), adopting a foreignizing translation strategy; whereas Unschuld focuses more on audience reception, interpreting TCM from a Western perspective and leaning towards a domesticating translation strategy. Therefore, this study selects Li's translation and Unschuld's translation (hereinafter referred to as Li's and Unschuld's versions) to compare their tendencies and differences in translation strategies and syntax with the AI-generated outputs, aiming to evaluate the readability and acceptability of the AI translations.

Example 3 前阴者，宗筋之所聚，太阴阳明之所合也。

Li's Translation: Qianyin (genitals) is the place where Zongjin (sinew connected with the genitals) gathers, Taiyin and Yangming [Channels] meet.

Unschuld's Translation: The front yin is where the basic sinews come together; it is where the major yin and the yang brilliance [conduits] merge.

ChatGPT Translation: The anterior yin (genital region) is where the ancestral sinews gather; it is also the meeting place of the Taiyin (Spleen) and Yangming (Stomach) channels.

DeepSeek Translation: The anterior yin is where the Ancestral Sinews converge and where the Taiyin (Spleen Channel) and Yangming (Stomach Channel) integrate.

From a strategic perspective, Li retained Pinyin (e.g., Qianyin/Zongjin) with parenthetical explanations of TCM concepts, adopting a foreignizing strategy.

Unschuld favored domestication, using common vocabulary (e.g., “前阴” -> “front yin”) and freely translating “Taiyin” as “major yin”. ChatGPT opted for a balanced approach: when translating “Taiyin” and “Yangming”, it retained the Pinyin “Taiyin/Yangming” while adding the Western medical terms “Spleen/Stomach” to lower the comprehension threshold, balancing cultural preservation with readability. DeepSeek primarily utilized a foreignizing approach to highlight academic rigor, retaining the Pinyin “Taiyin/Yangming”. Compared to Unschuld's colloquial tendency, DeepSeek employed words like “converge” and “integrate”, emphasizing the process of Qi and blood intersection embodied by “聚” (gather) and the characteristics of meridian “合” (integration) in TCM.

Syntactically, Li used a simple single-clause structure, while Unschuld used two coordinate clauses. This is similar to ChatGPT's structure, which connects two coordinate clauses with a semicolon, aligning better with English expression habits. DeepSeek used “and” to connect two “where” clauses, conceptually highlighting the dual positional attributes of “anterior yin”, resulting in greater logical rigor.

Lexically, Li employed “Pinyin + explanation” and mostly simple vocabulary (e.g., “合” -> “meet”). Unschuld utilized free translation based on personal interpretation, such as translating “Yangming” as “yang brilliance”. Both ChatGPT and DeepSeek opted for a literal translation of “Zongjin” (宗筋), rendering it as “ancestral sinews”. However, in TCM, “Zongjin” refers to the gathering of the sinews of the Three Yin and Three Yang meridians at the anterior yin area; it carries no connotation of “ancestors” or “lineage” as implied by the word “ancestral”. This exposes the limitations of AI tools in handling such culturally deficient concepts.

Example 4 厥阴之厥，则少腹肿痛，腹胀，泾溲不利，好卧屈膝，阴缩肿，飧内热。

Li's Translation: Jue [Syndrome] of Jueyin is characterized by lower abdominal swelling and pain, abdominal distension, unsmooth urination and defecation, preference to lie down with the

knees bent, swelling and shrinkage of the genitals and feverishness over the inner side of the legs.

Unschuld's Translation: Recession in the ceasing yin [conduit] results in [the following:] the lower abdomen is swollen and has pain, abdominal distension; the jing and the urine do not pass freely. [Patients] like to lie down with their knees bent. The yin (i.e., the sexual organ) is shrunk and swollen. The shins are hot at their inner side.

ChatGPT Translation: In Jueyin Jue, there is swelling and pain in the lower abdomen, bloating, difficult urination and defecation, a preference for lying curled with knees drawn up, retraction and swelling of the genitalia, and internal heat affecting the joints and bones.

DeepSeek Translation: for the Syncope syndrome of the Jueyin channel, it manifests as swelling and pain in the lower abdomen, abdominal distention, inhibited urination, preference for lying with knees bent, retracted and swollen genitals, and heat in the inner aspect of the thighs.

Strategically, Li continued his customary method of Pinyin annotation combined with parenthetical explanations for core concepts, while adopting relatively domesticated expressions for symptom descriptions to facilitate understanding. Unschuld tended to paraphrase core concepts, breaking the original long sentence into multiple short sentences for easier comprehension. In contrast, ChatGPT retained the Pinyin format of "Jueyin" while selecting terms commonly used in modern medicine, such as "bloating" and "difficult urination and defecation", to enhance clinical practicality and readability, leaning overall towards clinical application. DeepSeek leaned towards a foreignizing strategy, adding the annotation "(Terminal Yin)" for "Jueyin", and selecting expressions like "inhibited urination" that better align with the TCM pathological mechanism of Qi stagnation, emphasizing professional normative standards and a faithful transmission of TCM concepts.

Syntactically, Li used a single main-trunk long sentence to command all symptom descriptions; Unschuld broke the original long sentence into

multiple independent short sentences, showing obvious colloquial features. ChatGPT adopted a medical report-style single sentence structure, using "presents with..." as the main trunk, followed by a coordinate structure of symptoms separated by commas. The omission of conjunctions made the overall structure highly compact. DeepSeek constructed a more rigorous academic framework, using the "for... it manifests as..." structure to clearly emphasize the causal relationship between the syndrome and its manifesting symptom cluster. Noun phrases were used to list symptoms (e.g., "swelling and pain...", abdominal distention..."), and prepositional phrases like "preference for lying with knees bent" were more formal and academic than ChatGPT's "a preference for lying curled with knees drawn up".

Regarding specific terminology, ChatGPT and DeepSeek exhibited differences. For "阴缩肿" (Yin shrunk and swollen), both conveyed the basic meaning, but DeepSeek clearly emphasized that the genitals were simultaneously in two states by using the conjunction "and" in "retracted and swollen". For "髀内热" (heat in the inner aspect of the shins), ChatGPT translated it as "internal heat affecting the joints and bones". This translation risks ambiguity or inaccurate localization due to the broadness of "joints and bones". DeepSeek translated it as "heat in the inner aspect of the thighs" [Note: 髀 refers to the lower leg/shin, but the translation is evaluated based on the model's output], which more accurately reflects the symptom's location.

These two examples demonstrate that ChatGPT balances readability and cultural transmission, facilitating comprehension for non-professional readers and showing advantages in cross-cultural popularization and public acceptance. DeepSeek focuses more on terminology accuracy and professional expression, with rigorous and complex syntax suitable for academic research. Although the powerful natural language processing capabilities of LLMs certainly reduce the time cost of translation, human intervention remains necessary in the translation practice of TCM texts to correct

philosophically mishandled concepts by AI, ensuring the final translation meets practical needs.

4. Discussion

4.1 Corpus Preferences and Pre-Training Data Differences

By comparing the terminology processing capabilities of ChatGPT and DeepSeek in Chapter 45 of the Huangdi Neijing, this study reveals technical differences between the two types of models. ChatGPT, based on the Transformer architecture, is pre-trained on massive unsupervised corpora and can recognize, generate, and understand natural language (Jiao & Chen, 2023). Despite its inherent advantages in general corpus training, it cannot directly provide necessary domain-specific knowledge. When ChatGPT is used for terminology extraction and background knowledge retrieval in translation tasks, deviations may occur, affecting translation accuracy (Y. Wang & Zhang, 2024). In this study, ChatGPT generalized the concept of “Shouzu” (hands and feet) into Western anatomical terms, weakening the core logic of different meridians assigned to hands and feet in TCM’s Three Yin and Three Yang differentiation. Furthermore, its dynamic generation mechanism triggered logical deviations in terminology; for instance, translating “气因于中” (Qi originates from the middle) as “Qi is thus impaired within the center”, mistranslating “中” (Zhong jiao, middle burner, including spleen and stomach) as a geometric “center”, failing to convey the spatial medical view of TCM’s “Triple Burner” differentiation.

In contrast, DeepSeek-V3 employs Multi-Head Latent Attention and a Mixture-of-Experts (MoE) architecture, laying a theoretical framework for efficient model inference (Wenqi Wang et al., 2025). Relying on a terminology annotation system built upon a vertical TCM knowledge base, it significantly improves terminology translation accuracy. This advantage stems from the V3 model’s FP8 mixed-precision training corpus deeply integrating standards like the WHO International

Standard Terminologies on Traditional Chinese Medicine, demonstrating the technical value of AI tools in terminology translation. However, DeepSeek still falls short in handling complex philosophical concepts. For example, the brief eight-character source text “三阴俱逆，不得前后” (San yin ju ni, bu de qian hou) encapsulates the spatiotemporal philosophy of “Three Yin and Three Yang” in TCM, yet the translation merely processes it as “when Reversal affects all Three Yin channels, causing inability to pass urine or stool”, exhibiting a dual loss of philosophical connotation. First, the character “俱” (ju, all/together) implies the synchronicity and holism of the reversed Qi flow in the three Yin meridians, reflecting the TCM view of “Correspondence between Nature and Human” — that the three Yin meridians in the human body simultaneously lose balance, much like the three Yin energies in nature. The translation fails to capture this philosophical logic. Second, “前后” (front and back) in the Huangdi Neijing often refers to “urination and defecation”, but deeper down, it metaphorizes the Qi transformation movement of “ascending, descending, exiting, and entering”. Zhang Jingyue of the Ming Dynasty emphasized in Lei Jing that this refers to “clear Yang failing to ascend, and turbid Yin failing to descend”. Translating it literally as “inability to pass urine or stool” reduces a philosophical concept to a mere symptom description. This exposes the bottleneck of AI translation in transferring philosophical connotations: when faced with terms condensing core TCM thinking, AI tools can only describe simple symptoms but struggle to grasp the TCM philosophy behind the words.

4.2 Human-AI Collaborative Pathway

The application of artificial intelligence in translating TCM classics is fundamentally an integration of technological rationality and humanistic spirit. ChatGPT and DeepSeek, with their powerful dynamic context reconstruction capabilities, alleviate the translator’s burden of processing surface linguistic structures, highlighting the value of AI as an auxiliary

translation tool. However, the exertion of such technological efficacy must occur within a translator-led framework. Specifically, a “Generation-Proofreading-Optimization”

collaborative model can be constructed: AI first generates an initial translation; then, DeepSeek, which performs better in terminology translation, is used for terminological correction and to supplement basic missing source-culture information; finally, human translators conduct philosophical and cultural optimization to produce high-quality translations. Just as the character “厥” (Jue) in the Huangdi Neijing denotes both a pathological state and the philosophical connotation of disordered Qi flow, human translators need to add explanatory notes to such expressions, freeing machine-generated terms from symbolic constraints and showcasing the vitality of TCM thinking.

The profound significance of this human-AI collaboration lies in reshaping translation ethics: machine-assisted translation reconstructs the translator’s subjectivity; technology must not deprive humans of the right to interpret cultural genes. In the AI-assisted translation process, translators must rely on their professional knowledge and cultural sensitivity to review and revise AI-generated content (Huashu Wang & Liu, 2024). For example, regarding culturally deficient concepts like “Zongjin”, the translator must revise the AI output. The creative translation “confluence of reproductive sinews” highlights the dual meaning of this concept, encompassing both reproductive and meridian-sinew functions. This collaborative model will effectively avoid the conceptual flattening caused by AI tools’ lack of algorithmic compatibility with philosophical connotations.

Looking to the future, the cultivation of translation talents in the TCM field needs to respond to the demands of the AI era from three dimensions. First, strengthen translators’ interpretative training in TCM classics, especially the study of annotated literature such as *Lei Jing*, enabling translators to identify machine misinterpretations of critical TCM concepts. Second, integrate the critical application of

AI tools into the teaching system, training translators to guide AI outputs toward TCM thinking through prompt optimization. Finally, highlight the translator’s role as a cultural arbiter in this collaborative model, integrating translator subjectivity and creativity into the translation process to ensure the accurate transmission of the source text’s cultural connotations and emotional nuances, thereby manifesting the wisdom of TCM culture amidst the technological torrent.

5. Conclusion

This study compared the performance of ChatGPT and DeepSeek in translating characteristic terms in Chapter 45 (Jue Lun) of the Huangdi Neijing, answering the core research questions posed: ChatGPT, utilizing a general corpus, exhibits a stronger ability to handle varying contexts and produces more natural and fluent translations, but suffers from terminological redundancies and errors. DeepSeek, utilizing a specialized TCM knowledge base, significantly improves terminology translation accuracy, and its generated translations are more aligned with academic research requirements. However, both tools have limitations in expressing philosophical connotations and fail to fully convey the core philosophical thought of TCM. The study indicates that a single tool cannot simultaneously satisfy the three requirements of terminology accuracy, textual readability, and philosophical depth.

In the future, with the continuous optimization of LLM technology and the ongoing accumulation of vertical TCM corpora, AI tools are expected to enhance their expressive power regarding philosophical connotations while preserving TCM authenticity. Further exploring more intelligent human-AI collaboration models and cultivating translators’ critical AI application skills may achieve an optimal balance among accuracy, readability, and cultural depth, providing a more robust technological and talent guarantee for the international dissemination of TCM classics. In the process, the translator’s role extends beyond editing and proofreading to include strategy formulation,

workflow guidance, and quality control (Yu, 2025). The ultimate standard of AI translation depends not only on the model's inherent capacity limits but also heavily on the translator's deep understanding of its operational mechanisms and the precision of

prompt design. This profound integration of human-machine collaboration will determine the future height and breadth of the intelligent translation of TCM classics.

References

- Chen, B., Zhang, Z., Langrené, N., et al. (2023). Unleashing the potential of prompt engineering in large language models: A comprehensive review. arXiv preprint arXiv:2310.14735.
- Gong, Y., & He, Z. (2022). On translators' language awareness in classical text translation: A comparison of the English translation of zangxiang terms in Huangdi Neijing by Li Zhaoguo and Wen Shude. *Journal of Chengdu Aeronautic Polytechnic*, 38(02), 85–88.
- Jiang, H., & Wu, J. (2019). Standardization of English translation of TCM terminology. *Chinese Science & Technology Translation*, 32(01), 5–8.
- Jiao, J., & Chen, T. (2023). Empowering English teaching with large language models: Four scenarios. *Technology Enhanced Foreign Language Education*, (02), 12–17, 106.
- Li, Z. (2008). Concept, principles, and methods for international standardization of TCM terminology translation. *Chinese Translators Journal*, 29(04), 63–70, 96.
- Li, Z., & Liu, X. (2005). *Huangdi Neijing-Su Wen Chinese-English parallel edition 2*. World Book Publishing Company.
- Qin, H., & Zhou, X. (2024). Research on large language models and language comparison. *Foreign Language Teaching and Research*, 56(02), 163–176, 318.
- Sahoo, P., Singh, A. K., Saha, S., et al. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv preprint arXiv:2402.07927.
- Unschuld, P. U., & Tessenow, H. (2011). *Huang Di nei jing su wen: An annotated translation of Huang Di's inner classic-basic questions* (2 volumes). University of California Press.
- Wang, H., & Liu, S. (2024). Impact of large language models on translator subjectivity and strategies for mitigation. *Foreign Languages and Translation*, 31(04), 13–19, 98.
- Wang, H., & Ma, K. (2023). Comparative study of AI translation applications: A case of biomedical texts. *Chinese Science & Technology Translators Journal*, 36(03), 23–26.
- Wang, W., Guo, M., Yang, D., et al. (2025). Review on the development and applications of large language models. *Journal of Zhongyuan University of Technology*, 36(02), 1–8.
- Wang, W., Qu, Q. (2024). Knowledge construction comparison in the English translation of Huangdi Neijing from the perspective of knowledge translation studies. *Guiding Journal of Traditional Chinese Medicine and Pharmacy*, 30(12), 201–205.
- Wang, Y., & Zhang, Z. (2024). Concerns and mitigation of AI translation with ChatGPT. *Chinese Translators Journal*, 45(02), 95–102.
- Yu, J. (2025). Exploration of DeepSeek translation capabilities: Cases in literary and financial texts. *Chinese Translators Journal*, 46(03), 172–179.
- Zhang, X., Tang, S., Sang, N., et al. (2025). Comparative study of English translation of TCM terminology based on AI-generated content language models. *World Journal of Integrated Traditional and Western Medicine*, 20(06), 1255–1262.
- Zhu, Y., & Wang, Z. (2025). Terminology translation research under the background of large language models: Status, problems, and prospects. *China Terminology*, 27(03), 12–20.

Zong, C. (2020). Prospects of human language technology. *Communications of the Chinese Association for Artificial Intelligence*, 10(01), 1–5.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of Brilliance Publishing Limited and/or the editor(s). Brilliance Publishing Limited and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.