ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

The Evaluation of AI Translation Competence in Ceramic Texts through the Lens of Knowledge Translation Theory: A Study of ChatGPT, DeepSeek, and Other Tools

Yujiao Feng

Jingdezhen Ceramic University, Jingdezhen, China
18784748127@163.com

Abstract: *This study investigates the knowledge transfer capability and cultural adaptability of artificial intelligence (AI) translation tools in specialized ceramic texts, grounded in Knowledge Translation Theory. Utilizing textual descriptions of Ming Dynasty blue-and-white porcelain plates featuring Zhong Kui (a Daoist deity) motifs from The Art of the Chinese Potter from the Han Dynasty to the End of the Ming as corpus, we developed a multi-dimensional "Human Scoring + Automated Metrics" evaluation framework centered on terminology translation and culture-laden terms. A systematic comparative analysis was conducted across six AI translation tools: ChatGPT, DeepSeek, DeepL, Doubao, Youdao Translate, and Baidu Translate. Results reveal significant disparities among tools in key areas: terminology recognition, semantic reconstruction, and cultural recontextualization. While contemporary AI tools demonstrate preliminary competence in processing ceramic texts, notable deficiencies persist in restoring cultural metaphors and maintaining terminological consistency. This research contributes to: (1) expanding the application scope of AI in specialized non-general domain translation, (2) enriching methodological frameworks for translation quality assessment, and (3) advancing the global dissemination of ceramic cultural heritage.*

Keywords: *Ceramic Texts; AI Translation; Translation Quality Assessment (TQA); Knowledge Translation; Cross-Cultural Communication*

1. Introduction

1.1 Research Background

Artificial intelligence (AI) is advancing rapidly, with large language models (LLMs) becoming key drivers of technological and industrial change. The global release of ChatGPT sparked widespread adoption, while China's DeepSeek has quickly risen as a major competitor. In early 2024, DeepSeek launched its first LLM using a distinctive "low-cost breakthrough + open-source" model, challenging industry norms and expanding the functional possibilities of AI language applications. ChatGPT facilitates cross-cultural communication through multi-turn dialogue, and DeepSeek supports image-text intertranslation via multimodal architecture. These tools are increasingly applied in scenarios such as real-time subtitling and intelligent customer service, helping overcome modern communication barriers [1].

DeepSeek excels in Chinese language processing and semantic generation. Although not designed specifically for translation, its core language abilities and open-source nature make it a meaningful participant in AI translation. In a globalized world, AI translation tools act not only as linguistic converters but also as cultural mediators—especially for non-material cultural heritage (e.g., Jingdezhen ceramics, Peking Opera) and culture-specific concepts. They face persistent "untranslatability" challenges with ethnocultural texts, requiring not only linguistic accuracy but also contextual and cultural understanding. DeepSeek's strengths in semantic generation, style transfer, and emotional simulation open new pathways for translating culture-rich texts. Future systems could combine DeepSeek for Chinese semantics with tools like ChatGPT or DeepL for English output—a hybrid framework with promising potential to surpass traditional machine translation and achieve genuine cultural translatability.

This study systematically compares the translation performance of ceramic texts across major AI tools, including DeepSeek, ChatGPT, DeepL, Youdao Translate, Doubao, and Baidu Translate.

1.2 Theoretical Framework

The application of artificial intelligence (AI) translation to specialized ceramic texts transcends mere linguistic substitution; it constitutes a complex process of knowledge transfer, cultural mediation, and communicative function reconstruction. This study aims to scientifically evaluate AI tools' competence in processing ceramic texts by establishing a multidimensional theoretical framework grounded in Knowledge Translation Theory and integrated with established translation quality assessment (TQA) models.

Originally developed in medicine and life sciences to describe cross-contextual dissemination of specialized knowledge, Knowledge Translation Theory has expanded into technical communication and terminology translation research (Wang Hongyin, 2020). Its core premise posits that translating specialized texts requires not only conveying linguistic meaning but, more critically, achieving equivalent restructuring of knowledge architecture. Ceramic texts—characterized by high terminological density, strong cultural embeddedness, and syntactic complexity—align precisely with this theoretical domain. Within this framework, translation unfolds through three interdependent phases: 1. Decontextualization: Extracting key knowledge elements (e.g., technical terms, production processes, cultural referents) from source texts; 2. Semantic Reconstruction: Reorganizing conceptual relationships in the target language while preserving logical coherence and semantic equivalence; 3. Recontextualization: Embedding restructured knowledge into culturally appropriate target-language expressions.

Building upon this tripartite model, our research interrogates whether contemporary AI translation systems possess the capacity for holistic knowledge transfer. We specifically analyze AI performance across three critical dimensions: Ceramic terminology recognition accuracy, Cultural context restoration fidelity, Logical coherence maintenance.

1.3 Current Research Landscape: A Critical Review

Research in AI-driven translation has evolved from rule-based to data-driven approaches [2]. Early systems like Feng Zhiwei’s 1984 Chinese-English framework relied on linguistic rules. Statistical methods improved through phrase alignment, but remained limited in semantic modeling [3]. The introduction of neural machine translation (NMT) in 2014 enabled end-to-end learning, and the Transformer model (2017) greatly improved handling of long-range dependencies with its self-attention mechanism. Current studies focus on multimodal fusion and context-aware large language models (LLMs), such as GPT-4, which uses large corpora for dynamic adaptation [4].

In China, researchers have developed approaches suited to the Chinese language. Sun Maosong’s team incorporated historical phonology from Guangyun in Middle Ancient Chinese translation, and Baidu’s “Wenxin Translator” optimizes classical poetry translation with tonal adaptation [5]. Shared challenges include low-resource language adaptation, cultural term preservation, and ethical bias control [6]. Key limitations persist: (1) shallow cultural metaphor transfer, (2) high computational costs versus mobile deployment needs, and (3) narrow automated evaluation criteria overlooking fidelity, fluency, and elegance. These issues motivate work on lightweight models and human-AI collaboration [7].

Although AI translation excels in general domains and real-time use, it remains understudied in specialized fields like ceramics, philosophy, and ethnography. Unlike law, medicine, or finance—which have standardized terminology—cultural and knowledge-rich texts lack systematic research on AI-mediated knowledge transfer, semantic reconstruction, and pragmatic adaptation [8]. Most studies focus on surface-level metrics like accuracy, neglecting deeper cognitive and translatological perspectives on how knowledge is filtered or reconstructed between languages. AI handling of cultural metaphors, historical context, and specialized discourse remains empirically observed but lacks a robust evaluation framework [9].

Within Knowledge Translation Theory, this study evaluates AI systems’ performance on knowledge-rich texts—particularly in semantic coherence, cultural re-contextualization, and cognitive equivalence—to advance explanatory and communicative capabilities in intelligent translation.

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

2. Experimental Design and Methodology

2.1 Research Questions

This study addresses three core research questions: Whether significant differences exist in the performance of various AI translation tools when processing ceramic texts; Whether human evaluation and automated assessment models demonstrate consistency and correlation in quality judgment; Whether current AI systems possess the competence to undertake specialized ceramic text translation—and if not, where their primary limitations lie. To address these questions, this study constructs a triple-validation pathway integrating human evaluation, automated metrics, and theoretical analysis to explore the boundaries and potential of AI translation tools within a specialized domain.

2.2 Participant and System Selection

To ensure representativeness and validity of experimental data, selection was conducted at two levels: human participants and AI translation systems.

(1) Human Participants: A cohort of 24 translation Master's students was recruited. All participants possess formal training in translation theory and practice, with specific exposure to technical and culturally specialized translation.

(2) AI Translation Systems

As shown in Table 1, six mainstream and representative AI translation tools were selected, covering both general-purpose and specialized systems from international and domestic developers:

Table 1: Overview of Six Representative AI Translation Tools

Tool Name	Type	Developer	Key Characteristics
ChatGPT	General LLM	OpenAI	GPT-4 model; supports contextual understanding
DeepSeek	General LLM	Zhipu AI	Domestic advanced model; emphasizes specialized text capability
Doubao	Dialog-oriented AI	ByteDance	Supports multi-turn interactive translation
DeepL	Translation-specific	DeepL GmbH(Germany)	Excels in European languages; high-quality English output
Youdao Translate	Translation system	Translation system	Supported by multi-domain terminology databases
Baidu Translate	Translation system	Baidu Inc.	Frequently updated model deployed in search systems

2.3 Corpus Selection

The corpus for this study was sourced from the text of The Art of the Chinese Potter from the Han Dynasty to the End of the Ming. A representative sample—the entry for PLATE CLII—was selected for analysis. The original English description is as follows: Dish with slightly everted rim. Fine porcelain with well-painted design in a fine greyish blue. In the centre a land-scape scene with a martial figure brandishing a pike and two small figures who may be demons. A gourd-vine scroll forms the border. Outside are lotus sprays supporting the Eight Buddhist Emblems (pa chi Hsiang). Mark in a double ring, Ta Ming Wan Li nien chih, made in the Wan Li period of the great Ming dynasty (I573-I619). The subject appears to be a representation of Chung Ku'ei, the demon-queller, who corresponds in many ways to the more familiar Shoki of Japanese legend.

2.4 Evaluation Framework

(1) Human Evaluators

A panel of five experts was assembled to conduct human evaluation: Three translation studies faculty members from leading Chinese universities, each possessing extensive experience in translation pedagogy and assessment, provided evaluations based on linguistic quality and strategic decision-making. Two bilingual ceramics specialists, actively engaged in ceramic museology, education, and

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

international exchange, assessed terminological accuracy and cultural adaptation in the translations. This composition ensures a balanced assessment integrating translational and domain-specific perspectives.

(2) Automated Metric Selection and Rationale

To enhance objectivity and reproducibility, this study complements human evaluation with COMET (Crosslingual Optimized Metric for Evaluation of Translation), a neural-based metric widely recognized in machine translation evaluation for its high correlation with human judgment. COMET leverages crosslingual pretrained models (e.g., XLM-R) to encode the source sentence, machine-generated translation, and reference translation (if available), modeling their semantic relations to output a continuous quality score. The publicly available wmt20-comet-da model was employed. This version operates in a reference-free manner, directly assessing translation quality using source-text and candidate-translation pairs without relying on human-rated references. Scores typically range from -1 to 1, with higher values indicating superior translation quality. Its use provides a robust, scalable complement to manual evaluation [10].

2.5 Experimental Procedure

This study employed a mixed-methods design, structured into the following phases:

Phase 1: Corpus Preprocessing

The source text, selected from The Art of the Chinese Potter from the Han Dynasty to the End of the Ming, underwent systematic annotation. A bilingual ceramics expert produced a reference translation and annotated key terminology and cultural entities to achieve bilingual alignment. Both source and reference texts were converted into UTF-8 encoded .txt files, with segments uniformly separated by ||| delimiters.

Phase 2: Translation Generation

The preprocessed source text was input into six AI tools (see Table 2) under uniform parameters. The instruction: “Please translate the following text in the style of professional ceramic literature, maintaining terminological accuracy and adding cultural annotations where necessary” was applied. Post-editing functions were disabled to retain raw output. For human translations, 24 MA translation students independently translated the same text; a consensus version derived through cross-verification served as the benchmark.

Phase 3: Multidimensional Evaluation

For the human evaluation, five experts independently scored all translations using a 1–5 Likert scale based on three key dimensions: Linguistic Quality (evaluating grammatical accuracy, syntactic complexity, and fluency), Terminological Accuracy (assessing consistency with domain-specific standards), and Cultural Adaptability (judging the appropriateness in rendering culture-specific terms, such as whether to translate "Wanli" as "Wan Li" or with a reign-period annotation). Mean scores and standard deviations were then calculated from these expert ratings. Additionally, an automated evaluation was conducted by executing customized scripts to compute COMET scores for all machine-generated translations.

Phase 4: Data Analysis

Phase 4 encompassed a comprehensive data analysis process based on the outcomes of previous stages. The experimental workflow began with corpus preprocessing, where a 105-word source text underwent terminology annotation and alignment, resulting in an annotated bilingual corpus in TXT format. This corpus then served as the input for the translation generation phase, during which batch AI translation was performed alongside human translation, producing six sets of AI-generated outputs and 23 human reference translations. All these translations subsequently entered the multidimensional evaluation phase, where they were assessed through both human scoring and automated metrics, yielding a score matrix in CSV format. Finally, this score matrix was analyzed in the data analysis phase using statistical methods and qualitative examination. Specifically, a one-way ANOVA was conducted to compare performance differences among the six AI tools, with a significance threshold set at *p* < 0.05. Spearman’s rank correlation coefficient (ρ) was calculated to evaluate consistency between automated metrics and human scores. Additionally, a qualitative analysis of low-scoring samples was carried out to identify recurrent error patterns, such as term mistranslation and cultural omission.

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

3. Results and Analysis

A standardized translation instruction was input into all systems: "Please translate the following text in the style of professional ceramic literature, maintaining terminological accuracy and adding cultural annotations where necessary."

Reference Translation:

Plate CLII: Blue-and-white dish with figure design and gourd-vine border (c. 1600)

A porcelain dish with a slightly everted rim. The body is of fine quality, decorated in underglaze blue with a design painted in a fine greyish-blue tone. The centre depicts a landscape scene: a martial figure brandishing a pike, accompanied by two much smaller figures, probably demons. The rim is adorned with a gourd-vine scroll. The exterior is painted with lotus sprays supporting the Eight Buddhist Emblems (Ba Jixiang). The base is marked within a double ring with "Da Ming Wan Li nian zhi" (Made in the Wan Li reign of the great Ming dynasty). The subject appears to represent Zhong Kui, the demon queller. Zhong Kui corresponds in many ways to the more familiar figure of Shoki in Japanese legend.

3.1 Comparative Analysis of Terminological Accuracy

Terminological accuracy is a primary metric for assessing specialized translation quality. The table below details the rendering of key terms across different translations.

Table 2: Translation of key terminology by different AI tools and student translators

English Term	Reference Standard	ChatGPT	DeepSeek	DeepL	Doubao	Youdao	Baidu	Students
rim	A	A-	A-	C	B	C	C	B
in a fine greyish blue	A	B	C	C	B	C	C	B
scroll	A	B	B	F	B	F	F	A
sprays	A	A-	B	F	C	F	F	A-
Eight Buddhist Emblems	A	A	A	F	A (Annotated)	D	F	A
double ring	A	A	A	B	A	B	B	A
Chung K'uei	A	A (Annotated)	A	F	A (Annotated)	F	F	A

Accuracy Key: A = Correct; A- = Nearly correct (minor nuance loss); B = Acceptable but generalized; C = Imprecise/Partial meaning; D = Major error; F = Complete mistranslation or nonsense.

Annotation Note: "(Annotated)" indicates the tool provided additional cultural explanation beyond the core translation.

As evidenced in Table 2, ChatGPT, DeepSeek, Doubao, and the student translators produced terminology closest to the reference standard. In contrast, DeepL, Youdao, and Baidu Translate demonstrated significant mistranslations or over-generalization, particularly for terms like Chung K'uei and scroll, adversely affecting textual fidelity.

Regarding terminological precision, ChatGPT and DeepSeek performed well on specific shape terms (e.g., everted rim → piekou) and ornamentation patterns (e.g., scroll → chanzhiwen), achieving terminological coverage rates of 83% and 78%, respectively. General-purpose tools like DeepL and Baidu, lacking domain-specific adaptation, exhibited high error rates (~40%), such as translating scroll as juanzhou (scroll, as in a book). Student translations neared the reference standard in accuracy (92%), though minor imprecisions persisted (e.g., rendering everted rim vaguely as weitu [slightly convex]).

Significant disparities were observed in handling culture-loaded terms. Doubao provided detailed explanatory notes for Eight Buddhist Emblems (bajixiang), surpassing DeepSeek's mere transliteration. Cultural omission was severe in some cases; Baidu's transliteration of Chung K'uei as zhonggui (Zhong Gui – "Bell Turtle") completely erased the term's cultural meaning, resulting in a 100% error rate for this item. Cross-cultural association was also inconsistently handled: while ChatGPT and Doubao noted the parallel between Zhong Kui and the Japanese Shoki, this contextual link was absent from the student translations.

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

3.2 Comparative Performance of AI Translation Tools

Statistical analysis revealed significant disparities in the overall performance of the six AI tools (ANOVA, *p* = .003). ChatGPT and Doubao demonstrated superior performance, both exceeding a score of 4.0, and were most effective at replicating the specialized style and cultural context of ceramic literature. The COMET evaluation results showed a strong correlation with human scoring trends, further validating the model's efficacy in semantic judgment. ChatGPT achieved the highest scores in semantic matching and information retention, indicating a pronounced capability for handling complex contexts and terminological translation. In contrast, Baidu Translate and Youdao Translate consistently exhibited terminology recognition errors and syntactical disorganization, ranking lowest in the COMET-based assessment.

Table 3: Multidimensional Translation Evaluation Results for Different Tools and Student Translations

NO.	Tool	COMET	Term Coverage(TC)	Cultural Score(CS)	Mean Expert Score
1	Student Trans	0.88	0.95	9	8.7
2	DeepSeek	0.87	0.90	9	8.5
3	ChatGPT	0.86	0.88	8	8.3
4	Doubao	0.82	0.85	8	7.9
5	DeepL	0.78	0.65	5	6.5
6	Youdao Trans	0.75	0.60	4	6.2
7	Baidu Translate	0.72	0.55	3	5.9

To systematically evaluate the six AI tools' performance on ceramic culture texts, four metrics were employed: COMET score, Term Coverage (TC), Cultural Score (CS), and Mean Expert Score.

The COMET score, calculated using the COMET-22 model, reflects semantic similarity to the reference translation (max 1.00). For instance, ChatGPT scored 0.86 due to its accurate preservation of cultural semantic relationships (e.g., linking Zhong Kui and Shoki), whereas DeepL scored only 0.78, penalized for semantic deviations like translating scroll as *juanzhou* (scroll for writing).

Term Coverage (TC) assessed the accurate translation of 8 core ceramic terms (e.g., underglaze blue with a fine greyish-blue tone, gourd-vine scroll, Eight Buddhist Emblems). DeepSeek correctly translated 90% of terms, while Youdao Translate managed only 12%.

Cultural Score (CS) measured the depth of culture-laden term handling: 0 points for no processing (e.g., DeepL's transliteration "Zhong Guyi" for Zhong Kui), 1 point for partial annotation (e.g., ChatGPT identifying the cultural concept without elaboration), and 2 points for a complete explanation (e.g., Doubao providing clear annotations for Eight Buddhist Emblems). Scores were summed across all cultural terms.

The Mean Expert Score is the average from the human evaluation (detailed in Table 3), originally out of 17, converted to a 5-point scale (Score / 17 × 5) for easier cross-comparison in this table (e.g., a raw score of 13 converts to 3.82).

3.3 Consistency between Automated and Human Evaluation

Automated metrics demonstrated moderate to strong correlations with human expert scores (Table 4), confirming the validity of the hybrid evaluation framework.

Table 4: Spearman's Correlation (ρ) between Automated Metrics and Human Scores

Automated Metric	Human Score(ρ)	p-value
COMET Score	0.81	<0.001
Term Coverage(TC)	0.76	<0.001
Cultural Score(CS)	0.65	0.018

The COMET score and Term Coverage showed the highest correlation with human judgment ($\rho > 0.75$), underscoring that semantic matching and terminological accuracy are paramount for assessing translation quality in specialized domains. The relatively lower correlation for the Cultural Score ($\rho = 0.65$) stems from a focus difference: human evaluators prioritized the detail of cultural annotations (e.g., whether the Wan Li reign mark was annotated with corresponding Gregorian calendar years), whereas the automated metric primarily assessed the presence of the cultural entity.

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

3.4 Distribution of AI Translation Error Types

Table 5: Distribution of AI Translation Error Types with Representative Cases

Error Type	Frequency	Example Case	Tool Involved	Primary Cause
Term Mistranslation	48%	Source: scroll Error: juanzhou (scroll for writing) Ref: juanyewen (scroll pattern)	DeepL, Baidu, Youdao	Confusion of ornamentation terminology; lack of domain adaptation.
Cultural Omiaaion	35%	Source: Chung K'uei Error: zhonggui (Zhong Gui – "Bell Turtle") Ref: zhongxun (Demon Queller)	Youdao	Erroneous transliteration leading to complete loss of cultural signification.
Incomplete Annotation	12%	Source: Eight Buddhist Emblems Output: bajixiang (Eight Auspicious Symbols) No explanation of symbols (e.g., Wheel of Law, Treasure Umbrella)	ChatGPT, DeepSeek	Failure to elaborate on the meaning of cultural symbols.
Factual/ Chronological Deviation	5%	Source: Wan Li period (1573-1619) Error: 1573-1669nian (1573-1669) Ref: 1573-1620nian (1573-1620)	Baidu Translate	Incorrect reign dates, violating ceramic periodization standards.

4. Reflection and Discussion on AI's Knowledge Translation Competence

4.1 Challenges and the "Crisis of Translation" in AI Knowledge Translation

4.1.1 Knowledge Disembedding: Semantic Network Fragmentation and Tacit Knowledge Obliteration

AI's limitations in translating ceramic knowledge stem from a systemic failure in contextual re-embedding, as emphasized by Knowledge Translation Theory. This results not only in fragmented explicit knowledge but, more critically, in the dismantling of tacit knowledge and cultural-semantic structures.

For instance, while a reference translation of "fine greyish blue" as "underglaze blue with a greyish tone" captures technical nuances—such as iron oxide content and firing methods—AI tools like DeepL reduce it to "grey-blue pattern," stripping away material and historical context. This reflects a "surface encoding trap," where AI fails to activate cross-disciplinary knowledge spanning chemistry, art history, and technology.

Similarly, most AI systems struggle to construct cross-modal semantic networks. When translating culturally embedded terms like "Chung K'uei" to "Shoki," tools such as DeepL and Youdao either transliterate inaccurately or omit cultural associations entirely. This reveals an inability to integrate iconography, ritual function, and historical transmission—collapsing rich cultural networks into linear word chains and exemplifying a "can't see the forest for the trees" limitation in knowledge translation.

4.1.2 Knowledge Coordination Failure: The Disjunction of Terminology, Style, and Cognitive Frameworks

AI exhibits significant limitations in knowledge coordination, giving rise to pronounced systemic contradictions. While it can achieve local terminological accuracy, its overall cognitive framework is often inaccurate and chaotic.

The mistranslation of lotus sprays supporting the Eight Buddhist Emblems is a prime example. The standard translation strictly adheres to the hierarchical knowledge structure of ceramic ornamentation:

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

the term spray refers to a specific compositional technique, supporting denotes spatial logic, and Eight Buddhist Emblems connects to religious symbolism. In stark contrast, Baidu Translate's output lotus spray supports eight elephants crudely amalgamates terms from disparate disciplines—botany (spray), mechanics (supports), and zoology (elephants)—creating a semantic chaos that violates fundamental disciplinary common sense. This error exposes a core AI deficit: the inability to grasp the text's overarching cognitive-semantic framework, causing it to fall into a “lexical substitution trap” and lose all discriminative power over knowledge hierarchies.

A severe stylistic disjunction is also evident. For instance, DeepL translated the professional description well-painted design in a fine greyish blue into the more colloquial exquisite grey-blue pattern, resulting in a loss of the archaeological catalogue’s academic tenor. Empirical data confirms this, with specialized AI tools averaging a stylistic adaptability score of 4.0, compared to only 2.3 for general-purpose tools (Table 5), underscoring how the degree of domain knowledge internalization directly dictates translation quality.

4.1.3 Collapse of Cultural Paratextual Mechanisms: The Cession of Interpretive Power in a Translation Crisis

A grave challenge for AI translation involves the implicit cession of cultural interpretive authority. When Baidu mistranslated the name of the deity Zhong Kui as Zhong Gui (a homophone meaning “Bell Turtle”), it not only erased the figure's folkloric function as a demon queller but also reconstructed a distorted cultural image through the negative connotations associated with the turtle in certain contexts [11]. This violent decoding tramples the relationship between the signifier and the signified. Worse, this could be accepted by non-specialist readers, potentially fostering a cognitive fallacy of a “turtle-shaped Zhong Kui.”

While Doubao’s translation noted that “the Japanese Shoki originates from the Tang Dynasty,” it failed to connect this to the historical context of Tang tomb figurine traditions or Sino-Japanese Buddhist exchange, reducing the cross-cultural comparison to an empty label. This phenomenon aligns with the “collapse of narrative frameworks” in Bruner's theory of cultural cognition [12]. By severing the intrinsic link between cultural symbols and their historical praxis, AI generates rootless knowledge fragments. The monopolization of interpretive authority by technological platforms risks marginalizing the academic discourse of professional communities [13].

4.2 Reconstructing Knowledge Translation Theory: From Crisis to Turning Point

4.2.1 Pathways for Knowledge Re-embedding in Hybrid Intelligent Systems

The key to addressing the aforementioned challenges lies in constructing human-AI collaborative hybrid intelligent systems to propel knowledge translation beyond mechanical transfer towards deep contextual re-embedding.

The core of this system is a dynamic knowledge graph verification mechanism. After the AI produces an initial translation, it first performs concept matching against a specialized ceramic terminology database (e.g., matching the term for a spray-pattern motif). Subsequently, tacit knowledge from artisan oral traditions is injected (e.g., linking the specific greyish-blue tone to its corresponding chemical composition and firing parameters). Finally, domain experts audit the completeness of cultural annotations. This closed loop of “machine screening, human empowerment” has shown promise in experiments; for instance, the translation accuracy of a specific interlocking scroll pattern increased by 23% after integrating object images and spectral data.

4.2.2 Upholding Domain Ethics and Academic Discourse Power

Concurrently, stringent domain-specific translation ethics must be established to safeguard academic discourse power. This involves: Prohibiting AI from independently handling sensitive texts, such as those detailing glaze recipes or cultural heritage authentication, to prevent academic ethical risks stemming from technical misjudgment. The annotation of knowledge sources in AI translations should be mandated (e.g., “Terminology source: China Ceramic Terminology Standard, 2020 Edition”) to dismantle the erosion of knowledge authority by the “black box” of technology. Establishing a “Ceramic Knowledge Translation Alliance” – comprising archaeologists, master artisans, and translators – to collectively set AI training standards and develop open-source terminology databases (e.g., OpenCeramic Glossary) is crucial to break the knowledge monopoly of commercial platforms. Only by positioning AI as a “knowledge porter” rather than an “interpretive agent” can we mitigate the current translation crisis and rebuild a human-AI collaborative knowledge ethics framework [14].

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

5. Conclusion: The "Impossible Trinity" of Knowledge Translation and AI's Inherent Limitation

The application of AI in translating ceramic texts remains constrained by the "Impossible Trinity" of knowledge translation—efficiency, depth, and ethics. This trilemma stems from both algorithmic limitations in adapting to contextual knowledge and the inherent paradox between technological rationality and cultural profundity.

Currently, AI cannot achieve true "knowledge translation." While capable of manipulating terminological networks, it lacks insight into their generative mechanisms or cultural embeddedness. It simulates academic language while simultaneously deconstructing the historical, technical, and symbolic systems that validate knowledge.

Without ethical constraints, AI-mediated cross-cultural dissemination risks becoming mere symbolic consumption rather than genuine knowledge exchange. Ceramic translation transcends linguistic transfer to regenerate civilizational genealogies. Authentic translation must reveal qinghua's multidimensional essence—from cobalt's chemical properties to its socio-political symbolism.

This "thick translation" cannot be achieved by AI alone. However, guided by human knowledge systems, AI can become a collaborator in safeguarding civilizational heritage. Only through human-AI synergy can we preserve the capacity to comprehend technical glaze significance, interpret cultural narratives, and understand symbolic intelligence in decorative patterns.

Acknowledgement

This paper is supported by Jiangxi Graduate Education and Teaching Reform Project (JXYJG-2024-082).

References

[1] Kaplan, J., McCandlish, S., Henighan, T., et al. Scaling Laws for Neural Language Models [EB/OL]. (2020-01-23). https://arxiv.org/pdf/2001.08361.pdf

[2] Feng, Z. Disambiguation Function Testing of Machine Translation Systems[A]. In: Chinese Information Processing Society of China, Chinese Association for Artificial Intelligence (eds.). Advances in Machine Translation Research: Proceedings of the 2002 National Machine Translation Symposium[C]. Location: Publisher, 2002: 231-254. (in Chinese)

[3] Feng, Z., & Zhang, D. Machine Translation and Human Translation Complement Each Other[J]. Journal of Foreign Languages, 2022(6): 77-87. (in Chinese)

[4] García, I. Translating by Post-Editing: Is It the Way Forward?[J]. Machine Translation, 2011(3): 217-237.

[5] Zhang, X., & Zhang, W. A Review of the Technologization Tendency in Translation[J]. Foreign Languages Research, 2016(5): 88-91. (in Chinese)

[6] Wang, H., & Wang, X. Research on Translation Technology in the Age of Artificial Intelligence: Application Scenarios, Existing Problems, and Future Trends[J]. Foreign Language and Literature Studies, 2021(1): 9-17. (in Chinese)

[7] Goodwin, C., Campbell, E., Biondi, E., et al. Performance of ChatGPT and Google Translate for Pediatric Discharge Instruction Translation[J]. Pediatrics, 2024, 154(1): e2023064025.

[8] Feng, Z., & Zhang, D. GPT and Language Research[J]. Technology Enhanced Foreign Language Education, 2023(2): 3-11+105. (in Chinese)

[9] Shangshan Ruoshui. Why Domestic DeepSeek-V3 Outperforms GPT-4o[N]. Computer News, January 6, 2025. (in Chinese)

[10] Lourenço da Silva, I. M., Schmaltz, M., Alves, F., Pagano, A., Wong, D., Chao, L., Leal, A., Quaresma, P., & Garcia, C. Translating and Post-Editing in the Chinese-Portuguese Language Pair: Insights from an Exploratory Study of Key Logging and Eye Tracking[J]. Translation Spaces, 2015(1): 144-168.

[11] Guo, J. Ethical Risks and Regulatory Strategies of Generative AI Large Language Models: A Case Study of DeepSeek[J]. Journal of North China University of Water Resources and Electric Power (Social Science Edition), 2025(4): 1-10. (in Chinese)

[12] Feng, Z., & Zhang, D. Paradigm Shifts in Linguistic Knowledge Production in Computational Linguistics[J]. Contemporary Rhetoric, 2024(2): 23-44. (in Chinese)

[13] Marcuse, H. One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society[M].

ISSN 2616-5783 Vol.8, Issue 8: 118-127, DOI: 10.25236/AJHSS.2025.080818

London & New York: Routledge and Kegan Paul, 1964.
[14] Zhang, W., & Zhao, B. Has Generative Artificial Intelligence Ushered in a New Era of Machine Translation? A Quality Comparison Study and Implications for Translation Education[J]. Journal of Beijing International Studies University, 2024, 46(1): 83-98. (in Chinese)