

A Study on Collaborative Innovation Models for Cultural Classic Translation from a Digital Humanities Perspective: A Case Study of the Construction of a Multilingual Database for "Dream of the Red Chamber"

Yujie Xu*. Chunyang Lin

School of Foreign Languages, Huaiyin Institute of Technology, Jiangsu Province, 223000, China

Email: 29322387@qq.com

How to cite this paper: Xu, Y. J., & Lin, C. Y. (2025). A Study on Collaborative Innovation Models for Cultural Classic Translation from a Digital Humanities Perspective: A Case Study of the Construction of a Multilingual Database for "Dream of the Red Chamber". Educational and Humanities, 1(3), 1-8. ISSN Print: 3104-4352; ISSN Online: 3104-4360.

<https://doi.org/10.63313/EH.9017>

Published: 2025-10-20

Copyright © 2025 by author(s) and Erytis Publishing Limited.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Abstract

The advent of the digital humanities era has brought new perspectives to the study of cultural classic translation. This paper, using the multilingual data-base of "Dream of the Red Chamber" as an example, deeply analyzes collaborative innovation models for cultural classic translation from a digital humanities perspective. By integrating digital humanities technologies with the translation process, this paper finds that multilingual databases transcend temporal and spatial constraints, enable interdisciplinary learning, and promote the production of high-quality translations. The paper argues that collaborative innovation models in the digital humanities can promote the diversification of translation subjects, the diversification of translation processes, and the sharing of translation results, thereby providing new solutions for the international promotion of cultural classics. Furthermore, this paper points out that collaborative innovation models in the digital humanities face challenges in standardization and the dilemma of technological and cultural differences. These challenges can be addressed through the construction of a standard system, the improvement of technical capabilities, and the deepening of cross-cultural communication.

Keywords

Digital humanities; Cultural classic translation; "Dream of the Red Chamber"; Multilingual database; Collaborative innovation

1. Introduction

As one of the classics of Chinese culture, the quality of its translation impacts the global dissemination of Chinese culture. Traditional translation methods, constrained by manual efficiency, disciplinary boundaries, and cultural barriers, are in-

creasingly unable to meet the demands of cultural exchange in the digital age. Digital humanities, a discipline at the intersection of computer science and the humanities, employs methods such as text mining, natural language processing, and visualization. It offers a fresh perspective on the study of the translation of literary classics. By building a multilingual database of "Dream of the Red Chamber," we can not only digitally preserve the text but also explore the cultural implications of its translation through social network analysis and knowledge graph construction. This article, using the construction of a multilingual database of "Dream of the Red Chamber" as an example, explores collaborative research mechanisms for the translation of cultural classics within the context of digital humanities. We hope to contribute to the globalization of China's outstanding traditional culture.

2. Literature Review

The development of digital humanities scholarship has ushered in a paradigm shift in translation research. Hu Kaibao and other scholars have found that digital humanities, through numerical calculations of the language style and frequency of loaded words in translated works, has successfully uncovered the mechanisms underlying the application of two translation strategies. A comparative study based on a corpus of parallel Chinese and English texts of "Dream of the Red Chamber" reveals that for Buddhist concepts like "Taixu Huanjing," Huo's translation strategy involves localizing the text, using similar expressions in the target language, rendering it "Fairyland." This approach facilitates understanding and acceptance by foreign readers, but also compromises the unique philosophical concepts of the original work. Translator Yang, on the other hand, adopts a foreignization strategy, refraining from translating culturally loaded terms like "Ah Q" and "Redology," and providing detailed annotations after translation, allowing readers to more authentically perceive the original culture. This comparative research approach relies on the automated segmentation and annotation of massive corpora using digital humanities tools, which greatly enhances the efficiency and accuracy of translation research.

Translating cultural texts is a complex, systematic undertaking encompassing numerous disciplines, including linguistics, literature, history, and sociology. Traditional, single-disciplinary perspectives are ill-suited to the task of cultural transplantation in the translation of cultural texts. Collaborative innovation models are emerging at this time. This concept requires members of different disciplines to share resources and collaborate based on digital platforms. For example, the Chinese-English parallel corpus of "Dream of the Red Chamber" created by the Faculty of Education at Nanyang Technological University in Singapore brings together researchers from diverse fields, including translation specialists, computer scientists, and cultural experts, to conduct comprehensive digital processing, from alignment to translation technique exploration. This brings new insights and approaches to the

translation of cultural classics.

Currently, numerous multilingual databases of "Dream of the Red Chamber" have been established both domestically and internationally. In China, the Chinese-English corpus of "Dream of the Red Chamber" developed by the National Social Science Research Group of Yanshan University uses "one-to-one" or "one-to-two" sentence alignment techniques to effectively facilitate comparative studies of translation styles across different translators. Overseas, the "Dream of the Red Chamber" digitization project at the Harvard University Library utilizes GIS technology to study spatial layout and interpersonal relationships in the novel. However, most multilingual databases of "Dream of the Red Chamber" are built around a single language or technical focus, making it difficult to cross-compare multiple language versions and technical methods. This inability to meet the increasingly diversified and comprehensive development needs of translated cultural classics.

3. Research Methodology

3.1. Case Selection

The selection of the multilingual database for "Dream of the Red Chamber" as a research case study was a deliberate decision. First, "Dream of the Red Chamber" is a shining jewel among the classics of Chinese culture. Its rich cultural connotations, unique artistic characteristics, and complex linguistic system make it highly representative and exemplary for translation research. Studying its translation is equivalent to gaining a deep understanding of the phenomena and issues that arise in the cross-linguistic and cultural transmission of this cultural classic. Second, the database is massive. It includes over 10 language versions, including the bilingual versions of Hawkes and Yang Xianyi, the German translation by Kuhn, the French translation of the complete works of Dormont, and the Chinese-Russian bilingual version from the Lenin Library of the Russian Academy of Sciences. This database encompasses nearly all major languages of Europe and Asia. The database's coverage of multilingual and cross-linguistic translation research is extensive and comprehensive. Finally, the database utilizes XML markup language and the TEI encoding standard. This allows for highly precise and accurate annotation of textual information, as well as the storage and retrieval of multimodal multimedia information such as images and audio. This improves the efficiency of data utilization and the breadth and depth of translation research.

3.2. Data Collection

The data sources for this study are quite extensive. First, original information was obtained through public databases, such as the "Chinese Classics Ancient Books Library" of the National Library of China and the European Digital Library (Euro-

peana). These databases contain a large number of ancient texts and digital resources, providing reliable information sources for research. Second, we fully utilized research data processed by scholars. For example, the annotated corpus in the Chinese-English Parallel Corpus of "Dream of the Red Chamber" is meticulously annotated by relevant scholars and has a high degree of accuracy. Third, we used natural language processing techniques to obtain structured data. For example, similarity matrices based on "doc2vec" can provide quantitative data support for research.

3.3. Analytical Framework

The collaborative innovation model in the digital humanities constructs a three-dimensional interpretive framework. At the technical level, emphasis is placed on text analysis, social relationship analysis, and visual synthesis. These three methods, combined, can fully explore the implicit information in the text, form complex relationship networks, and present research results through visualization. At the personnel level, cross-disciplinary collaboration is emphasized, fostering communication between translation scholars, technical personnel, and cultural researchers, leveraging different disciplines to learn from each other's strengths and weaknesses, and injecting new thinking and ideas into the research process. At the resource collaboration level, we prioritize the integration of multilingual corpora, annotation materials, and research methods, breaking down resource barriers and achieving efficient resource utilization to further support research progress.

4. Research Findings

4.1. Technical Collaboration: From Single Tool to Complex System

The multilingual database of "Dream of the Red Chamber" systematically advances translation research through the organic integration of various digital humanities technologies. For the Chinese and English corpora, during the text data mining and alignment process, the "edit distance" algorithm is used to automatically align Chinese and English sentences, effectively addressing the time-consuming and low-accuracy issues of manual alignment. Detailed records in the database show the results of sentence alignment in the translated passage "Jia Baoyu's Anti-Imperial Examination Thoughts." For example, comparing Hawkes's translation with Yang Xianyi's, the automatic alignment achieved an accuracy of 92%. This data demonstrates that the algorithm performs relatively well for this type of general sentence alignment. However, the accuracy of cultural word alignment is only 68%. This significant discrepancy reminds us of the importance of technology's sensitivity to culture and highlights key issues we should be mindful of when using technology. The character relationship map created using Gephi on social networks subtly reveals the influence of translation strategies on character portrayal. For example, Hawkes's translation of "Wang Xifeng" (王熙凤) as "Hsi-feng Wang" downplays or even ig-

noses the deeper cultural connotations of the surname. Yang Xianyi, on the other hand, retains "Xifeng Wang" and explains the significance of the surname "Wang" in Chinese families in a footnote, allowing readers to fully understand the character's inner meaning.

A heat map of translation strategy percentages generated using Tableau software visually illustrates the different approaches taken by various translations to culturally loaded terms. The database chart shows that Yang Xianyi's translation of "Buddhist terminology" has the highest proportion of foreignization strategies, reaching 73%, aiming to maximize cultural similarities and differences for the target language readers. Hawkes's translation, at 31%, is more accommodating to the target language audience's level of acceptance.

4.2. Collaboration: From Individual Research to Teamwork

Corpus construction requires interdisciplinary teamwork. Translation experts play a crucial role in this effort, meticulously annotating the corpus and researching translation strategies. For example, regarding the culturally rich expression "Taixu Huanjing," translation theory experts provide detailed interpretations based on its cultural significance for future research.

Technical staff develop automatic alignment tools and visualization platforms. They write Python scripts that can simultaneously batch process data in multiple languages, significantly reducing data processing time and improving efficiency. Cultural scholars provide guidance on the project's historical background and cultural knowledge, such as explaining the "imperial examination system" in the text, enabling team members to understand the underlying cultural logic of the text.

Teams use GitHub to share code and data, continuously improving the database through an agile development approach. Real-world data shows that this interdisciplinary approach has increased translation research efficiency by 40% and reduced errors by 25%, demonstrating the effectiveness of group collaboration.

4.3. Resource Collaboration: From a Closed System to an Open Ecosystem

The database opens an API interface, enabling resource sharing among research institutions worldwide. In terms of corpus co-construction, we completed a link with the Chinese-English parallel corpus of the Center for Chinese Foreign Language Education Research at Beijing Foreign Studies University, effectively adding German, French, and other languages to the database's corpus resources.

Tool reuse involves the use of the "Translation Analysis Toolkit" developed by Nanyang Technological University, Singapore, which can monitor translation quality in multiple languages and greatly supports the expansion of the database's functionality. User participation is achieved through crowdsourcing. Readers around the world can mark difficult translation sections, such as crowdsourcing annotations for

cross-cultural analysis of "Redology" terminology. This pools global wisdom and opinions, enhancing the usability and accuracy of the database.

The resource integration effect is remarkable, with the database corpus size leaping from 100,000 to 5 million, and the language coverage expanding from 2 to 15, creating an open and diverse research ecosystem.

5. Discussion

5.1. Advantages of the Collaborative Innovation Model

The collaborative innovation research approach has many benefits for translation studies. In terms of optimizing translation quality, the digitization of linguistic features in translation corpora can effectively illuminate the translator's cultural attitudes. For example, the database found that Yang's translation of "Anti-Feudal Consciousness" was 65% literal, while Hawkes's was 28%, clearly reflecting different cultural perspectives.

In promoting cultural interaction, multilingual databases provide a good platform for cross-cultural exchange for international academia. For example, French scholars, based on the database, discovered that Kuhn's romantic approach to "Descriptions of Love" resonates with German literary traditions, enabling cross-cultural learning.

In terms of teaching innovation, the database is suitable for "flipped classroom" teaching methods, allowing students to use visualization features to independently reflect and research translation methods. Actual data shows that classes using the database saw an average increase of 15 points in translation proficiency test scores, actively promoting reform and innovation in teaching methods.

5.2. Challenges

However, the collaborative innovation model also has its limitations, one of which is the issue of data standardization. Different databases use different indexing methods, making integration between them impossible. For example, the compatibility of the TEI encoding standard with DublinCore metadata requires the construction of a data standard model. Regarding technical ethics, AI devices can ignore cultural taboos. For example, machine translation software can easily misinterpret concepts like "yin and yang." Therefore, we need a human review mechanism to address this issue, ensuring that machine translations are more accurate and better integrated into local culture.

In terms of cross-cultural cognition, international users often have cultural biases when marking translation difficulties. For example, Western users often interpret the "imperial examination system" negatively. Technology must be neutralized by drawing on diverse cultural contexts to more fully and accurately reflect cultural behavior.

5.3. Response Strategies

To address these issues, proactive measures should be adopted. Establish standardized standards, such as the "Specifications for the Construction of a Multilingual Database of Cultural Classics," to unify standards and technical formats, laying the foundation for the integration and sharing of data resources.

Strengthen technical ethics regulations, establish a translation quality evaluation system, and add cultural adaptability as an evaluation criterion to ensure that automated products incorporate cultural factors into their machine translation.

Further strengthen cross-cultural exchanges. Leverage international cooperation projects, such as the "Belt and Road" cultural classics translation initiative, to enable collaborative research between Chinese and foreign scholars, jointly advancing the pace of classics translation research.

6. Conclusion

The collaborative innovation model for translating cultural classics within the context of digital humanities possesses vibrant vitality and creativity. Leveraging technological collaboration, it integrates numerous digital humanities technologies, transforming classic translation from a single translation research approach into a systematic one. Relying on collaborative efforts, it overcomes disciplinary barriers, bringing together interdisciplinary scholars such as translation scholars, engineers, and cultural studies scholars, leveraging their strengths to enable cross-disciplinary research. Leveraging resource synergy, it breaks down closed research resources and builds an open research ecosystem, enabling the open circulation of resources.

The model used, for example, in the construction of a multilingual database for "Dream of the Red Chamber" has proven highly effective. It effectively improves translation quality, allowing translated texts to better reflect the meaning of the source language, effectively promotes cultural exchange, enables diverse cultural groups to better understand Chinese culture, supports teaching innovation, and expands engaging teaching content.

However, this model presents challenges with data standards, ethics, and cross-cultural communication. We should establish data standards, strengthen ethical awareness, and enhance cross-cultural communication. Future research will explore the selection of AI for translation and the powerful role of blockchain in protecting data copyright, enabling cultural classics to be better disseminated globally.

Reference

- [1] Franco Moretti. Conjecture on World Literature[J]. New Left Review, 2000(01).
- [2] DING Yuan. A Study on Dietetic Words Translation of Hong Lou Meng Based on Functional Equivalence Theory[D]. Lanzhou: Lanzhou Jiaotong University, 2017.
- [3] YANG Xiaoru. A Comparative Study on the Translation of Cuisine Names in Hong Lou Meng from the Perspective of Dietary Culture[D]. Xi'an: Shaanxi Normal University, 2011.

- [4] Nida, Eugene A. Toward a Science of Translation[M]. Shanghai: Shanghai Foreign Language Education Press, 2003.
- [5] YANG Dingyi. A Functional Equivalence Perspective to the Study of English Versions of Food Names Translation in Hongloulou[D]. Chengdu: Southwest Jiaotong University, 2012.
- [6] CHEN Zhiming. Tofu and Related Products in Chinese Foodways[J]. The World
- [7] Venuti, Lawrence. The Translator's Invisibility: A History of Translation[M]. New York: Routledge, 1995: 4.