Academic Journal of Management and Social Sciences
ISSN: 2958-4396 | Vol. 6, No. 1, 2024

Domestic Research on Computer-Assisted Translation: Frontier Hot Spots and Evolutionary Paths

Nan Ma

Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract: In recent years, computer-assisted translation is one of the important teaching contents in MTI majors, and the discussion of its content helps to cultivate students' translation ability. Based on the bibliometric theory and CiteSpace analysis tool, this paper systematically combed the domestic computer-assisted translation literature in the past ten years, and analyzed the knowledge mapping from the aspects of keyword co-occurrence, keyword clustering, keyword mutation, etc., in order to explore the frontier hotspots of the research in this field and its evolution path. The results show that computer-assisted translation has significant development prospects, and its research content covers a wide range and tends to be diverse. Based on the research results, this paper also makes an outlook on the further development of this field.

Keywords: Computer-Assisted Translation; CiteSpace; Visualization and Analysis; Current State of Research.

1. Introduction

Under the background of globalization, the demand for translation in China's language service industry is increasing year by year, and the market of this industry has put forward higher requirements for translation quality and efficiency, i.e., translators in the era of artificial intelligence should not only have basic translation skills, but also be proficient in computer-aided translation tools such as SDL Trados and memo Q. Different from traditional machine translation, Computer-Aided Translation (hereinafter referred to as CAT) is a translation activity centered on translators, who assisted in translation by the core technology of translation memory. With CAT tool, translators improve their translation efficiency and save time[1]. Translators can also use CAT tools to check the quality of their own translations, so as to have a better control of the text. From cultivating traditional literary translators to cultivating interdisciplinary talents in the field of "language plus", the cultivation program for MTI students in domestic universities has been gradually adjusted, so MTI students will face more technical learning and practice in their courses. Based on the bibliometric tools, this paper will summarize the current situation and discuss the hot spots, development trends and shortcomings of the current research in this field. And on the basis of this paper, the article will make a prospect of the research in this field, in order to provide a reference for the academic field.

2. Research Design

(1) Data sources

The data of this paper was obtained from the database of CNKI, and the search subject was "computer-aided translation". The time span was 2012-2022, and the search was carried out in three directions, namely, the main topic, the secondary topic, and the title of the article. After removing the irrelevant literature, the database finally obtained 500 related research literature. All literature includes academic journals, thesis and dissertation, and conference literature. Academic journals bring innovation, knowledge and academic communication[2]. Thesis and dissertation emphasize originality and theory, and have high reference value for the research field of this paper. Conference literature has real content and detailed knowledge, and is a vital way to obtain academic information[3].

(2) Research methodology

In this paper, the 500 selected literature were exported in Refworks format and then analyzed with the help of CiteSpace 6.1.R6 (64-bit) software. The time span was set as 2012-2022, the time slice was set as 1 year, and the clipping method was chosen as path simplified networks (pruning sliced networks), and the main attentions were on the keyword co-occurrence map, the keyword mutation analysis, and so on. From the keyword co-occurrence and keyword clustering, the hotspots of research can be dignified; from the keyword mutation analysis and keyword timeline, the trends and evolution paths of research can be valued.

(3) Research tools

CiteSpace is a bibliometric software for citation visualization developed by Professor Chaomei Chen at Drexel University, USA[4]. The software presents the structure, regularity and distribution of scientific knowledge by means of systematic visualization, and supports a variety of bibliometric studies for researching the research frontiers, the number of publications, and keywords in a particular field.

3. Data Analysis

(1) Keyword co-occurrence map

The resulting keyword co-occurrence map of the research literature related to "computer-assisted translation" was carried out by CiteSpace, and the keyword co-occurrence map of computer-assisted translation in Figure 1 was obtained. The size of the keyword nodes in the graph is proportional to the frequency of the word, and the larger the colored keyword nodes are, the higher the frequency of the word. After sorting and categorizing, the top 10 keywords in terms of keyword frequency are listed

Table 1. Frequency of keywords in computer-assisted translation

No.	Word	Number of occurrences	Mediation centrality	Year of issuance
1	machine translation	72	0.38	2012
2	post-editing	29	0.03	2018
3	translation technology	28	0.19	2013
4	translation teaching	26	0.12	2012
5	term base	20	0.16	2013
6	translation memory	18	0.03	2012
7	translation	16	0.11	2012
8	terminology management	13	0.02	2013
9	artificial intelligence (AI)	13	0.07	2018
10	SDL Trados	8	0.03	2015

(see Table 1), which are machine translation, post-editing, translation technology, translation teaching, term base, translation memory, translation, terminology management, artificial intelligence and SDL Trados, respectively. The results show that the keywords with the highest frequency of occurrence are machine translation (72), post-editing (29) and translation technology (28). Most of the high-frequency words are related to computer translation technology, and the probability is that they are the hot content of computer-assisted translation research. What's more, mediation centrality has great research value. A strong mediation centrality means that the keyword is closely related to other keywords, indicating that it has a high potential in future research. If the value of mediation centrality is greater than 0.10, it indicates that the term occupies importance in the research field. According to Table 1, the keywords with mediation centrality greater than 0.10 are ranked in descending order: machine translation, translation technology, term base, translation teaching, and translation. These five words have a higher chance of co-occurrence and closer interconnection in the knowledge network of the field of computer-assisted translation, that is, they are more influential in the subject area.

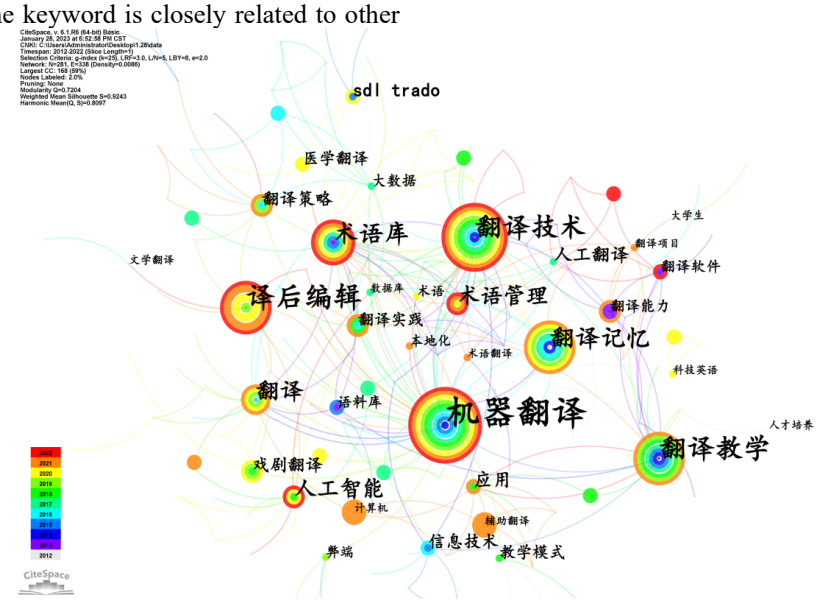

Fig 1. Computer-aided translation keyword co-occurrence map

The frequency of keywords in this research field is shown in Table 1, and the year of occurrence indicates when the keyword first appeared. The results show that the earliest date of the literature in the field of "machine translation" can be traced back to 2012. Experiments related to "machine translation" first appeared in the 1950s. Due to the limitations of the technical conditions of the times, the translations obtained by machine translation were poorly readable and had a poor reading experience[5]. It was not until 1990 that the research and application of machine translation gradually became familiar to people. After more than half a century, machine translation has been able to meet the user experience of many scenarios. In 2005, Google launched its online translation service and created a statistical-based machine translation system[6]. Subsequently, translation software has been supported by more advanced technology, and a variety of online translation software not only supports mutual translation of multiple languages, but also realizes the functions of speech translation, graph translation and real-time translation. The use of cloud technology in machine translation has penetrated into people's daily work and life; in 2012, according to the requirements of China National Committee for Graduate Education of Translation and Interpreting, the teaching concepts and methods of translation were further promoted, and "translation teaching" and "translation" also became influential in the field; the application of "translation memory" can be traced back to the 1970s. According to Bowker, translation memory is similar to a kind of memory bank that records the content of repeated terminology translation. In translation activities, translation

memory can effectively reduce the activities of repeating translation contents and improve translation efficiency[7]. Translators can add or delete the contents of translation memory, so that they do not need to translate twice when encountering repeated contents, which greatly shortens the translation time; then in 2013, the keywords "translation technology", "term base" and "terminology management" began to enter the field of computer-aided translation. Traditional translation methods are no longer suitable for dealing with huge amounts of texts, and the industry follows the innovations in the field of translation and uses relevant translation technologies to improve the quality of translators' work and save enterprise costs. Yuan Yining defines "translation technology" as specific information technology and electronic tools used by the translation industry in the translation, including term base, software localization, machine translation and so on[8]. A "term base" is a linguistic database created to efficiently manage a huge amount of terminology, integrating the functions of collecting, storing, modifying and categorizing, aiming at solving the problem of text inconsistency in the translation, as well as facilitating the accuracy of the translation; Keyword "terminology management" in Table 1 refers to the activities carried out on terminological resources in a specific textual environment, which generally include the collection, processing, storage, compilation, organization and other practical elements of terminology. Effective terminology management ensures consistency in the application of terminology in all parts of a text and removes barriers between writing and translation[9].

According to an Internet survey conducted by Hou Xiaochen and Yan Yuzhuo, a computer-aided translation program at Peking University in 2008, about 42% of the translators would choose to use the German SDL Trados software in their work. Compared with CAT tools such as Wordfast and YiCAT, Trados has more comprehensive functions and modularized parts, supports multiple file formats, and is better in matching function, terminology management, multiple editing environments and so on[10]. However, recent studies have found that Trados is not as easy as it should be. In recent years, research has found that the preparation part required by Trados consumes too much time. Problems such as insufficient matching rate of the term base have gradually emerged, so the subsequent improvement of Trados-related functions has also been included in the scope of key considerations; in 2018, the field of computer-assisted translation began to focus on "post-editing" and "artificial intelligence". Thanks to the rapid development of technology, various artificial intelligence can roughly meet the translation requirements of the normal language level, but for the fixed content and high requirements of the translated text, the translators gradually use the mode of human-computer interaction translation to complete the translation activities, i.e., first use the machine translation to process the text, and then use the translation technology software to carry out secondary processing of the text to fully improve and edit the management of the translated text. Then the translation efficiency can be greatly improved.

(2) Development of researches in the field of computer-assisted translation

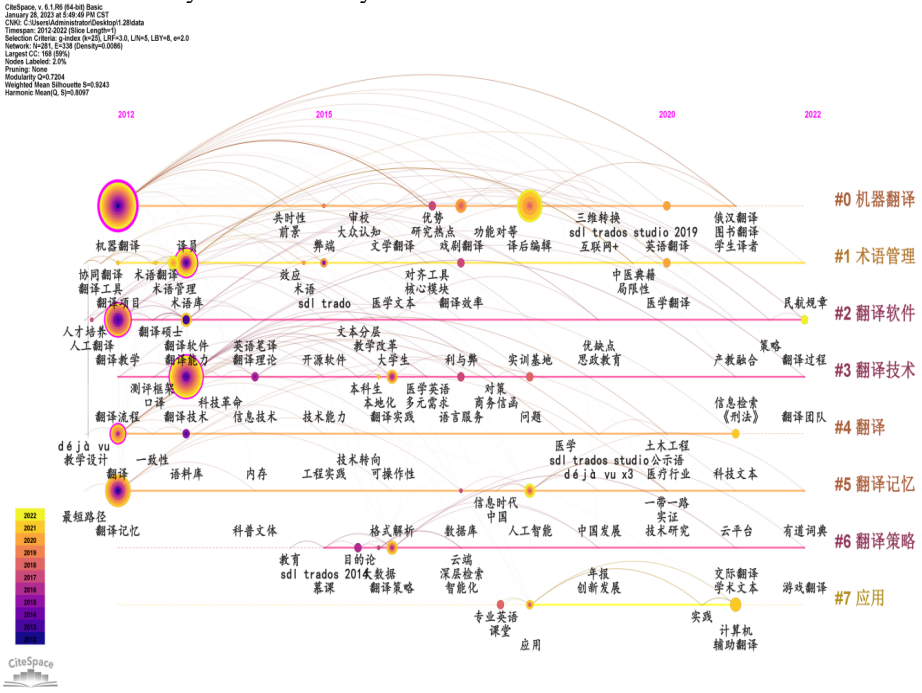

Fig 2. The timeline chart of keywords clustering

The relationship between keyword clustering and year can be visualized more visually using the literature timeline chart. The horizontal coordinate represents the time when the keywords first appeared (Figure 2), and the research topics are arranged vertically in descending order according to the size of the clusters. As can be seen from the figure, as early as 2012, the high-frequency keywords in this field appeared, and the research literature centered around "machine translation", "translation teaching", "translation", "translation memory" and other clustered hotspots. Since 2013, keywords such as term base, translation software, translation capability, corpus, and translation technology have appeared in computer-aided translation research, indicating that the translation field has begun to pay attention to the use of translation technology, and at the same time, the various translation processes have put forward higher operational requirements for translators, i.e., they need to parse, organize, reconstruct and integrate the translated content through the acquisition of linguistic

knowledge[10]. SDL Trados became a hot keyword in the field of computer-aided translation in 2015, after which Trados appeared many times in terminology management and translation strategy, becoming a key learning technology in the translation community and setting a new operational standard for translation technology in the field. From 2016, translation practice and translation strategy began to appear as keywords, and MTI teaching also followed the hotspot and put the teaching focus on translation practice content as well. Then in 2018, the translation community began to focus on post-editing, artificial intelligence and related applications. Until 2022, translation software and translation technology still receive attention from the academic community, and all realize that CAT tools are indispensable for promoting globalization and localization processes [11].

As can be seen from Figure 2, the development of computer-assisted translation research can be divided into three stages. The first stage is 2012-2013, which is the high-speed development stage of computer-assisted translation research, and machine translation has become the hotspot of research in this time period. Translation technology spreads rapidly, and the academic community also conducts in-depth research in translation teaching, terminology management, translation technology, etc.; the second stage is 2014-2018, which is a period of rapid development in the field, and the research field covers a wide range of aspects, including SDL Trados, translation practice, translation field, post-editing, artificial intelligence, etc.; the third stage is 2019 to the present. It continues the second phase (2014-2018) on computer-assisted translation research hotspots, and presents a multidisciplinary dispersed research. Most of the research content is combined with the contemporary popular computer content, and the direction of development is more clear, and the main related words are medical translation and informative texts.

4. Deficiencies and Prospects of Domestic Computer-Assisted Translation Research

First of all, some MTI colleges are still lacking professional teaching faculty of computer-assisted translation courses. Since computer-assisted translation is a course that has been developed in China only in the last decade or so, most teachers are traditional language learners, and there are few lecturers who get the knowledge of computer and language. Although traditional language teachers have strong language potential, they may have problems such as insufficient concern for the development of the industry in the CAT field. They lack operational ability in the long-term teaching. Therefore, the professional teaching faculty is particularly important. To address the shortage of professional teachers in this field, the following two aspects can be considered: MTI institutions can arrange as much as possible for "computer technology + language " talents to teach and improve the quality of teaching course; in the exchange of institutions, colleges with advantageous teaching faculty can play a leading role in promoting exchanges between universities, so as to reach a better teaching efficiency.

What's more, there are still deficiencies in the cultivation of knowledgable talents in this field. To carry on the Chinese spirit in international communication, it is necessary to tell stories about China well. China is eager to make its voice heard. Telling a good Chinese story requires a higher level of integrated talents, and the market demand for such talents has always been on the high side. For universities, how to promote the cultivation of compound talents in the field of computer-assisted translation is an issue worth thinking about. It is also worth paying attention to how colleges should combine language knowledge with different disciplines, such as information technology, law, journalism and communication, architecture, materials, etc., so as to cultivate integrated talents who can adapt to the new era.

5. Conclusion

The article conducted a visualization and analysis of 500 literature on computer-assisted translation during 2012-2022. The data was measured in terms of keyword co-occurrence map and research hotspots. It is found that (1) the research related to computer-assisted translation is developing gradually, and the number of journal articles is growing fast, but there is a downward trend during 2020-2022, and the development is slightly slow. The current research disciplinary background of this field is relatively small, and it has not yet been able to become a multidisciplinary cross-study field. (2) among the hotspots of computer-aided translation research, machine translation and post-editing are still highly concerned topics in the field. Through keyword clustering analysis, it is found that the research hotspots of computer-aided translation mainly focus on the three aspects of post-editing, translation technology and term base, and there is a certain literature distribution in each aspect. (3) corpus, information technology, and medical editing are the keywords in the research field while the term post-editing has the highest mutation intensity, indicating that it occupies a higher position in the research field of computer-assisted translation. Overall, in the current background, corpus, information technology, medical editing and post-editing will maintain the heat and become the key content of computer-assisted translation research. However, there are still some deficiencies in the field of computer-assisted translation in the cultivation of professional teaching faculty and composite talents, teaching method and school-enterprise cooperation. In the future, we should strengthen the exchange and cooperation between institutions, promote the progress of translation technology, accelerate the cultivation of composite talents adapted to the requirements of the new era, and optimize the cooperation between schools and enterprises, so that we can know the development trend of the industry, and improve the translation efficiency more quickly, and make full use of the potential advantages of computer-assisted translation.

References

[1] Huashu Wang, Defeng Li, Liqing Li. Research on the teaching of translation technology in Master of Translation (MTI): problems and countermeasures[J]. Computer-Assisted Foreign Language Education in China, 2018, (03):76-82+94.

[2] Xiaoping Liu. The Special Role of Academic Journals in Knowledge Innovation[J]. Journal of Editing, 2001, (03):135-137.

[3] Guo Hua. On Important Scientific and Technological Information Resources--Conference Literature[J]. Library Work and Study, 2006, (01): 25-27.

[4] Chen Yue, Chaomei Chen, Zeyuan Liu, Zhigang Hu, Xianwen Wang. Methodological functions of CiteSpace knowledge graph[J]. Research in Science, 2015, 33(02): 242-253.

[5] Li Xin. An Analysis of the Cultivation Mode of Master of Translation--Taking Three Universities in Asia as an Example [J]. Beijing Education (Higher Education), 2022(08): 61-64.

[6] Wang Feng, Shiyu Peng, Yan Dan. A study on assisting Chinese to English translation based on Google's new function--Taking the English translation of university web pages as an example[J]. Shanghai Journal of Translators, 2010, (04): 66-69.

[7] Bowker Lynne. Computer-Aided Translation Technology: A Practical Introduction [M].University of Ottawa Press: 2002-01-01.

[8] Yining Yuan. Translation technology and the training of Technical Translation Talents in China [J]. Chinese Science & Technology Translators Journal, 2005, (01): 51-54.

[9] Huashu Wang, Zhang Zheng. Research on terminology management in translation projects[J]. Shanghai Journal of Translators, 2014, (04): 64-69.

[10] Yubin Zhu, Xiaoqian Chen. A comparative study of four common computer-aided translation software at home and abroad [J]. Computer-Assisted Foreign Language Education in China, 2013, (01): 69-75.

[11] Duoxiu Qian. Reflections on Computer-Assisted Translation Teaching [J]. Chinese Translators, 2009, 30(04): 49-53+95.