

Translation Technology in the Language Services Industry: Concepts and Classification

Li Kefu

Liaoning University of International Business and Economics, Dalian, Liaoning, 116052, China

Abstract: This paper systematically reviews the concepts, development, and classification of translation technology in the language services industry. Translation technology has evolved from Computer-Assisted Translation (CAT) to modern Machine Translation (MT) and now covers the entire process of pre-translation, during-translation, and post-translation. The paper details the core technologies of each stage: pre-translation stage, during-translation stage and post-translation stage. With the development of AI technology, translation technology has been upgraded from tool assistance to a "cognitive collaboration" paradigm, but the dominant role of humans remains irreplaceable.

Keywords: Translation technology; CAT; MT; Language services industry

DOI: 10.62639/ssperi06.20250109

1. Introduction

In recent years, with the development of artificial intelligence technology, especially the emergence of ChatGPT, many scholars in translation studies have once again focused their research on the field of translation technology. People are amazed by the excellent performance of generative artificial intelligence tools like ChatGPT in translation tasks. Some researchers "even advocate the theory of technological threat and technological substitution, causing a certain degree of technological anxiety and technological fear among translation practitioners and learners" (Wang Huashu, Wang Xin). This phenomenon also reflects the emphasis that translation practitioners place on translation technology. The 2025 China Translation Industry Development Report points out that in the language services industry, "the use of translation technology among translation companies and practitioners has further expanded. In 2024, the proportion of companies using translation technology to complete translation tasks reached 97.1%, an increase of 2.5% compared to 2023; the proportion of practitioners using translation technology to complete translation tasks reached 97.7%"(69). What is translation technology? What technologies does translation technology include? The author, in combination with existing research results, further combs through the concepts and classification of translation technology.

2. The Development and Concept of Translation Technology

(1) The development of translation technology

The history of translation technology is relatively short, but it has developed rapidly. The failure of machine translation in the 1960s led to the emergence of Computer-Assisted Translation (CAT). From the beginning of Computer-Assisted Translation in 1967 to 2013, the development of translation technology has gone through four periods: the first stage, the budding stage of translation technology (1967-1983); the second stage, the steady growth stage (1984-1992); the third stage, the rapid development stage (1993-2003); the fourth stage, the global development stage (2004-2013) (Chan Sin-wai).

(Manuscript NO.: ERI-25-9-113001)

About the Author

Li Kefu (1978-), Male, Manchu, born in Jinzhou, Liaoning. An associate Professor at the English Department of Liaoning University of International Business and Economics. Main research areas are English language teaching, second language acquisition and translation.

In the first stage, truly commercialized Computer-Assisted Translation systems did not emerge, and translation technology failed to impact translation and the translation industry. In the second stage, Computer-Assisted Translation companies represented by Trados and Star Group began to appear. The release of Trados in 1992 marked the entry of Computer-Assisted Translation systems into the commercial age. In the third stage, more commercial translation software emerged, with more built-in functions, support for more file formats, and support for more language translations. This period was also the time when the Windows operating system dominated, and Trados was the leader in the translation market. In the fourth stage, translation technology showed a global development trend, with different countries successively developing different translation assistance software. The main features included: first, compatibility with the Windows operating system and Microsoft's Office suite. Second, workflow control integrated with CAT systems. Third, the popularity of networking and online systems. Fourth, the adoption of new formats. Since 2013, the field of translation technology has undergone tremendous changes. After more than seventy years of development, translation technology has become an umbrella concept that covers four fields: machine translation, Computer-Assisted Translation, localization, and voice translation (Chan Sin-wai).

(2) The concept of translation technology

The definition of translation technology can be considered from both narrow and broad perspectives. In a broad sense, translation technology can generally refer to all the technologies used by translators from pre-translation to during-translation, and until post-translation to complete translation tasks. Lynne Bowker, a Canadian linguist and professor of translation studies at the University of Ottawa, believes that translation technology refers to the various technologies used in human translation, machine translation, and Computer-Assisted Translation, including the use of basic computer tools as well as specialized translation tools. Wang Huashu redefined translation technology: "Translation technology refers to the various technologies comprehensively applied by translation service personnel in the translation process, including pre-translation technologies such as format conversion, resource extraction, word count, repetition rate analysis, task analysis, terminology extraction, extraction of repeated segments, and pre-translation technology; during-translation technologies such as assisted spelling, assisted input, electronic dictionary and parallel corpus query and verification, translation memory matching, and terminology recognition; and post-translation technologies such as quality checking, translation format conversion, post-translation typesetting, translation product language testing, and language asset management" (Wang Huashu, Wang Xin). The definitions of the above two scholars can both be regarded as broad definitions.

The language services industry also understands translation technology from a broad perspective. For example, the Chinese Translators Association's Written Translation Service Standard states that "translation technology includes a series of various auxiliary tools used by translators, bilingual proofreaders, monolingual proofreaders, and other personnel". The scope of translation technology in this standard is also relatively broad, including Content Management Systems (CMS), writing systems, desktop publishing systems, word processing software, Translation Management Systems (TMS), Translation Memory (TM) tools, Computer-Assisted Translation (CAT), quality assurance tools, revision tools, localization tools, Machine Translation (MT), terminology management systems, project management software, speech-to-text recognition software, and other existing and future related technologies. The 2025 China Translation Industry Development Report points out that "the use of translation technology among translation companies and practitioners has further expanded. In 2024, the proportion of companies using translation technology to complete translation tasks reached 97.1%, an increase of 2.5% compared to 2023; the proportion of practitioners using translation technology to complete translation tasks reached 97.7%. Among various translation technologies, the technologies with higher usage rates among translation companies and practitioners are collaborative translation platforms, translation memory tools, and online dictionary tools" (69). The report also lists the specific translation technologies used by translation companies and practitioners, including collaborative translation platforms, translation memory tools, online dictionary tools, terminology management tools, large corpora, automated engineering and technical tools, bilingual sentence libraries, automatic QA tools, speech recognition tools, terminology prompt tools, natural language processing tools, and 11 other categories.

From a narrow perspective, translation technology should only refer to the various Translation Management Systems

(TMS), Translation Memory (TM) tools, and Computer-Assisted Translation (CAT) technologies used by translators in the process of translating the source language into the target language. The main purpose of these technologies is to improve translation work efficiency. Xu Bin, Qian Duoxiu, Zhang Xiaojun, and others also hold similar views, believing that CAT technology has both narrow and broad senses. In a broad sense, CAT covers all software and hardware technologies that can provide auxiliary support in the translation process, and these electronic tools can greatly improve the efficiency of language and translation cultural exchanges; in a narrow sense, CAT refers to specialized translation assistance software designed to improve translation efficiency and optimize the translation process (Wang Huashu, Wang Xin).

3. Classification of Translation Technology

(1) The process of translation services

The International Organization for Standardization (ISO) has released a written translation project standard, "Translation services—Requirements for translation services (ISO 17100:2015)," which divides the translation process into three production stages: pre-translation, during-translation, and post-translation. The Translators Association of China has adopted the written translation project standard released by the International Organization for Standardization (ISO) and released a group standard for the translation industry in 2016: "Translation Services: Requirements for Written Translation Services (T/TAC1-2016)." In this standard, the written translation project process is divided into three interrelated processes: pre-production process and activities, production processes, and post-production processes (Cui Qiliang). Essentially, these are different expressions of pre-translation, during-translation, and post-translation.

The language services industry further refines the translation process. The pre-translation process includes four items: First, the translation service provider responds to the client's translation inquiry and conducts a feasibility analysis; Second, the translation service provider quotes a price; Third, the translation service provider signs a translation agreement with the client; Fourth, the translation service provider processes client information related to the project. The during-translation process includes six items: written translation, self-checking, bilingual proofreading, monolingual proofreading, copyediting, and inspection and delivery. It should be noted that monolingual proofreading and copyediting are not mandatory processes. The post-translation process includes two items: First, client feedback; Second, project closure management (Cui Qiliang).

(2) Overview of pre-translation technology

Wang Huashu believes that pre-translation technology includes format conversion, resource extraction, word count, repetition rate analysis, task analysis, terminology extraction, extraction of repeated segments, and pre-translation technology (Wang Huashu, Wang Xin). Alan Kenneth Melby believes that pre-translation technology includes alternative terminology extraction, terminology research, new text segmentation, and bilingual text alignment and indexing of the original text. In summary, in the pre-translation phase, the technologies and tools applied mainly focus on improving efficiency, ensuring terminological consistency, pre-processing text, and providing optimized support for subsequent translation or Machine Translation (MT). They can be divided into the following six categories: text preprocessing technology, terminology and memory bank preparation technology, Machine Translation pre-processing technology, project analysis and resource matching technology, AI-assisted pre-translation technology, and pre-quality inspection technology.

1) Text preprocessing technology

Text preprocessing technology includes techniques for format standardization, file format conversion, and text segmentation, with file format conversion being the primary focus. The main purpose of format standardization is to clean up garbled characters, unify punctuation marks, and remove redundant spaces or hidden characters, etc. Commonly used tools include regular expressions and scripting tools. File format conversion is a commonly used technique in text preprocessing. As the name suggests, file format conversion tools can convert one format to another while keeping the file

content unchanged (in an ideal state). Typically, this involves converting files from an uneditable state to an editable state, such as converting files from PDF or image formats to doc or docx. Common tools include Pandoc, Adobe Acrobat DC, and OCR tools (such as ABBYY FineReader). Text segmentation involves dividing a continuous piece of text into the smallest segments that can be independently processed according to "translation units." Common text segmentation tools include LangChain TextSplitters, TextSplit, NLTK, and the Spacy library.

2) Terminology and memory bank preparation technology

Terminology and memory bank preparation tools include tools for terminology extraction and translation memory alignment. Terminology extraction refers to the process of automatically identifying and extracting terms with specific domain meanings, high frequency of occurrence, and the need for unified translation from monolingual or bilingual texts using manual or automated means during translation or knowledge engineering. Translation memory alignment refers to the process of breaking down existing "source + target" files that are translations of each other (usually in Word, PDF, HTML, TTX formats) into the smallest translation units (sentence segments) and pairing them one-to-one to generate translation memory files in formats such as TMX, TMX.gz, or SDLTLM that can be imported into CAT tools. Tools like LF Aligner and SDL Trados WinAlign use sentence-level alignment algorithms to pair existing bilingual texts into TM entries.

3) Machine translation preprocessing technology

Common machine translation preprocessing includes source language enhancement and controlled language. Source language enhancement eliminates noise or errors in the original text through spelling correction tools, grammar correction tools, and traditional-to-simplified Chinese conversion tools, reducing mistranslations caused by "dirty input" in MT. For example, by using automatic spelling correction, grammar checking, auto-completion, and disambiguation prompts, issues such as spelling errors, grammatical errors, missing words, and ambiguous words in the original text can be addressed. Common tools include LanguageTool, Microsoft Editor, and custom regular expression scripts. Controlled language, on the other hand, uses tools like Acrolinx and Congree to enforce simplified syntax or terminology lists, optimizing the quality of MT output. Essentially, controlled language imposes limitations on vocabulary, syntax, and style at the writing end, making the original text naturally suitable for MT.

4) Project analysis and resource matching technology

Project Analysis and Resource Matching is a quantification-driven decision-making process. It involves a baseline assessment of the project scope, complexity, and risks through systematic indicators (Project Analysis). The assessment results are then mapped to the available resource pool (people, translation memories, terminology databases, MT engines, budget) to achieve optimal configuration and meet constraints (Resource Matching). This process follows the principles of the "Resource Management Plan" and "Estimate Activity Resources" knowledge areas in the PMBOK and is completed through the CAT/PM toolchain. For example, tools like Trados Studio and Xbench are used for word count and repetition rate analysis, and TM match analysis (100%/fuzzy match) is used to estimate workload and cost. Tools like ViStyle and PerfectIt are used for style guide and rule checks, with predefined formatting rules for the target language.

5) AI-assisted pre-translation

From a process perspective, pre-translation is a feed-forward meta-operation aimed at the cross-cultural functional reconstruction of the source text. It generates a situation model that can guide the selection of translation strategies by systematically describing intra-textual variables (e.g., lexical density, syntactic complexity, genre markers) and extra-textual variables (e.g., author's intent, receiver's expectations, historical-cultural context). From a product perspective, pre-translation is an automated draft generation mechanism. It uses translation memories (TM), terminology bases (TB), and machine translation engines (MT) to perform batch segment alignment and replacement on the source text, outputting a raw target-language hypothesis without post-editing. This reduces subsequent cognitive load and increases the reuse rate of language assets. With the development of AI technology in recent years, pre-translation can be performed using AI tools. Common pre-translation technologies include Adaptive MT and Terminology Injection. Common tools include DeepL Pro

and Google AutoML Translation, which can fine-tune the MT engine in real time based on the client's TM. Common tools for terminology injection include SDL Trados Studio and MemoQ.

6) Quality pre-inspection

Quality Pre-inspection refers to a set of activities that involve preventive and systematic checks and assessments of elements that may affect quality before a product, project, or process is officially implemented (or delivered, produced). The core goal is to identify risks in advance and intercept defects to ensure that subsequent stages meet the established quality requirements, thereby reducing rework, lowering costs, and improving overall quality performance. In the language services industry, QA tools are used to check for terminology consistency, tag completeness, and number matching before files enter the CAT tool to identify potential errors in advance. For example, Verifika and QA Distiller are used for translatability checks to detect non-translatable content (such as text embedded in images, unextracted variables). Content moderation tools (such as Azure Content Moderator) are used for cultural sensitivity scans to mark cultural symbols that may need localization.

(3) Overview of translation technology in the during-translation phase

The during-translation phase refers to the time window from when the translator opens the file to be translated and formally begins "producing the translation" until the first draft of all segments is completed. The core demands of this phase are to reduce cognitive load, increase translation speed, and ensure terminological and stylistic consistency. During this phase, translators can adopt the "machine translation + post-editing" translation mode, interactive machine translation mode, or use AI tools to assist in polishing and reviewing the translation (Cui Qiliang).

Translation technologies in the during-translation phase include assisted spelling, assisted input, electronic dictionary and parallel corpus query and verification, translation memory matching, and terminology recognition. Melby believes that during-translation technologies include automatic terminology lookup, translation memory lookup, and machine translation (Wang Huashu, Wang Xin). Nowadays, with the development of technology, during-translation technologies have evolved from "standalone CAT" to a real-time human-computer coupling system that integrates five dimensions: TM, TB, MT, LLM, ASR, and QA. Translators can simultaneously access five types of resources—memory, terminology, models, voice, and quality inspection—within the interface, achieving a dynamic balance between efficiency, quality, and artistry.

Translation memory and segment reuse technologies, including real-time translation memory lookup (TM Live Lookup), sub-segment matching, and cloud TM synchronization, can minimize repetitive work and enhance terminological and expressive consistency during the translation process. These technologies use algorithms, data structures, and interfaces to segment, index, calculate similarity, and recall existing bilingual corpora in real time.

Terminology consistency technologies, including real-time terminology base pop-up, adaptive terminology injection, and large language model terminology alignment, ensure that the same term maintains an unambiguous and stable mapping between the source and target languages in multi-document, multi-translator, and multi-iteration environments. These technologies rely on algorithms, rules, and interfaces to achieve consistency.

Machine translation and adaptive technologies enable the general base model to quickly and controllably adapt to specific domains, users, or real-time feedback without retraining all parameters. This results in improved translation quality and stylistic consistency for target scenarios.

Large language model collaboration during translation forms a dynamic closed-loop interaction between the large language model (LLM) and the translator or computer-assisted translation (CAT) system within the real-time workflow of "translation in progress." This collaboration leverages context awareness, knowledge enhancement, human-computer feedback, and multi-model scheduling to achieve a collaborative paradigm of "learning while translating and optimizing while translating."

Real-time quality assurance uses sensors, the Internet of Things, artificial intelligence, and edge computing to

instantaneously collect, analyze, and provide feedback on key quality parameters during the execution of production, service, or translation processes. This enables early detection and correction of defects, preventing batch errors and continuously optimizing the process.

Multimodal in-translation collaboration allows translators to simultaneously utilize various modalities such as text, images, audio, and video during real-time translation. This is achieved through a collaborative framework that integrates large language models (LLM), visual-language models (VLM), and small specialized models (SLM). This technology enables cross-modal disambiguation, emotional restoration, cultural context completion, and dynamic quality optimization.

(4) Overview of translation technology in the post-translation phase

The post-translation phase is the final checkpoint "before delivery" in the translation process. Its core task is to systematically optimize, inspect, restore formatting, and accumulate knowledge of the already generated translations to ensure that the final deliverables meet client requirements in terms of language, format, functionality, and compliance. Post-translation technologies include quality checking, translation format conversion, post-translation typesetting, language testing of translation products, and language asset management (Translation Technology Tutorial, 3). Melby categorizes terminology consistency checks, non-permitted terminology checks, missing segment checks, and formatting and grammar checks as post-translation technologies.

Specifically, post-translation technologies include translation quality enhancement techniques such as automatic post-editing (APE), context consistency scanning, style and sentiment alignment, and machine QA technologies. They also include format and layout restoration techniques such as format restoration, text expansion/contraction compensation, and protection of non-translatable strings. Terminology and memory bank sedimentation technologies include incremental alignment and entry into the database, terminology consistency auditing, and terminology frequency-weight analysis. Compliance and functionality testing technologies include regulatory compliance scanning, which can automatically compare regulatory terminology and prohibited words.

4. Conclusion

In summary, translation technology has evolved from a "tool-assisted" model to a "cognitive collaboration" paradigm. Memory, terminology, machine translation, and large language models form a closed loop before, during, and after translation. Data, algorithms, and human experience are coupled in real time, redefining the role of translators and the logic of quality control. The scope outlined in this paper shows that technology is no longer just a lever for efficiency but has become a "context generator" and "decision agent." Its boundaries have expanded from monolingual to multimodal and from text to knowledge networks. Despite the powerful advantages demonstrated by technology, we must not overlook the value of humans as the subjects of technology. Margaret Boden, a British researcher in the philosophy of artificial intelligence, points out that language, creativity, and emotion still pose significant challenges to artificial intelligence (Li Hanji). Learning technology, utilizing it, and becoming a master of it is a necessary path for every translator.

References

- [1] Wang Huashu, Wang Xin Research on Translation Technology in the Age of Artificial Intelligence: Application Scenarios, Existing Problems, and Future Trends [J]. *Foreign Languages and Literatures*, 2021 (01): 9-17
- [2] Xing Yutang 2025 China Translation Industry Development Report [R]. Beijing: China Translators Association, 2025.
- [3] Chan Sin-wai *Routledge Encyclopedia of Translation Technology* [M]. Routledge, 2023:1-31. [4] Cui Qiliang Research on the Application of Artificial Intelligence Technology in Translation Projects [J]. *New Perspectives in Translation Studies*, 2023 (2): 41-51.
- [5] Li Hanji The Evolution and Reconstruction of the Relationship between Translation Technology and Translators in the Age of Artificial Intelligence [J]. *Journal of China West Normal University*, 2022 (05): 113-118.
- [6] Yu Jing, Liu Kanglong Reshaping Translation Studies: Paradigm Shift and Future Directions under the Influence of AI Technology [J]. *Foreign Languages*, 2024 (07): 72-81.