

技术文档的机器翻译及译后编辑

——以LangChain官方技术文档为例

陈明芳, 陈乙

武汉工程大学外语学院, 湖北 武汉

收稿日期: 2025年11月17日; 录用日期: 2025年12月19日; 发布日期: 2026年1月4日

摘要

译后编辑作为提升翻译质量的关键环节, 能有效弥补机器翻译的不足。本文以LangChain官方技术文档及其谷歌翻译结果为研究对象, 结合技术文档的特点与翻译要求, 探究技术文档机器翻译的常见问题及译后编辑策略, 解决机器翻译存在的术语错译、句式生硬、语篇衔接松散等问题, 提高用户体验, 并为技术文档的精准化、规范化翻译提供一些参考。

关键词

技术文档, 机器翻译, 译后编辑

Machine Translation and Post-Editing of Technical Documents

—A Case Study of LangChain's Official Technical Documents

Mingfang Chen, Yi Chen

School of Foreign Languages, Wuhan Institute of Technology, Wuhan Hubei

Received: November 17, 2025; accepted: December 19, 2025; published: January 4, 2026

Abstract

As a crucial process in improving translation quality, post-editing can effectively offset the drawbacks of machine translation. Integrating the characteristics and translation requirements of technical documents, this paper takes LangChain's official technical document and its Google Translate results as research subjects and defects some common machine translation problems, such as mis-translated terminology, unnatural sentence structures, and loose textual coherence, and explores

post-editing strategies for technical documents from the three dimensions respectively. The purpose is to enhance user experience and provide some references for the accurate and standardized translation of technical documents.

Keywords

Technical Documents, Machine Translation, Post-Editing

Copyright © 2026 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

当前,机器翻译已成为翻译过程中的重要手段,其每天由机器翻译产生的译文字词数已超过人类翻译的所有译文语料的总和[1]。但机器译文目前还很难做到一步到位,必须有译后编辑这个环节。译后编辑因结合了机器翻译的效率和人工翻译的精准度,为全球化交流提供了高质量的语言支持,而成为提高翻译质量的必须步骤[2]。

技术文档是包含产品详细信息和操作指南的科技文本,广泛应用于计算机、工程、医学等行业[3],主要为用户提供产品简介、操作说明或维修指南等信息,是产品和服务不可或缺的技术说明文件。技术文档翻译关系企业全球化进程,是语言服务行业的重要内容。而现实中,技术文档的翻译仍有很多值得进步的空间[4]。

LangChain 是美国的一个大语言模型应用开发框架,其技术文档对于使用者来说不可或缺。本文以谷歌翻译作为机器翻译的平台,摘选 LangChain 技术文档作为翻译素材,需说明的是本研究为英文软件开发文档汉译,且采用单一平台,在此基础上总结该机器翻译在该领域的不足之处,以减少人们阅读技术文档的语言障碍,推动技术全球化和知识共享。

2. 技术文档的特点和翻译要求

了解技术文档的语言、术语、风格等方面的特点,是对技术文档进行翻译并保证翻译质量的前提。

2.1. 技术文档的特征

技术文档属于普通科技语体,具有五个方面的显著特征。(1) 行文严谨:技术文档的核心功能是传递准确信息,因此行文必须避免任何可能引发误解的表述,呈现出正式庄重的语体色彩。(2) 内容详尽且简洁:在确保内容详尽完备的同时,技术文档也应避免冗余,力求表述简洁,以提高阅读效率。(3) 可读性强:技术文档应使用简单易懂的语言,确保读者能快速理解核心信息,尤其是对非母语者的阅读者更为友好。同时还需考虑翻译需求,便于跨语言传播[5]。(4) 运用专业术语:专业术语对精准描述技术概念和组件属性等具有不可替代的作用;使用的术语需符合规范,避免自造术语或误用。(5) 表述整体统一:整份文档的术语使用、格式规范及语言风格应保持一致,有助于建立文档的整体性,更能凸显其结构的清晰、条理化,提高文档的易用性。

2.2. 技术文档的翻译要求

技术文档的翻译并非简单的语言转换,需要精准传递技术信息,适配目标读者需求,基于技术文档

的语言特征, 其具体翻译要求可梳理为如下五个方面: (1) 准确性: 译文必须精确反映原文的专业术语和概念, 确保无错译、漏译, 是技术文档翻译的核心前提。(2) 可读性: 译文应通顺明白, 使读者能够轻松理解, 避免复杂冗余的表述。如果原文存在信息架构不合理、措辞不当、内容冗余等不足, 可以进行优化。(3) 规范性: 译文应遵循目标语言的技术文本语体规范, 力求语言规范、准确一致, 避免语言生硬和难以理解, 便于目标语言读者查找使用。(4) 专业性: 译者应具备相应领域的技术背景或专业知识, 以确保能正确理解原文, 并实现译文在技术层面的严谨性与可信度。(5) 一致性: 专业术语、缩略语、语言风格、格式规范等在整份文档中应保持统一, 不仅能降低读者的认知成本, 避免因表述差异引发困惑, 更能凸显文档的专业性。

3. 机器翻译的应用和译后编辑

随着技术的进步, 机器翻译对于科技文本的翻译有明显进步, 但仍和人工翻译有差距, 主要原因是翻译系统更倾向于学习和使用训练语料中高频出现的语言结构, 而容易忽视低频结构[6]。机器翻译译文缺乏准确性、可读性及连贯性, 容易出现低级误译和漏译, 有极大的文化缺陷, 无法解决多义词、歧义词及语境带来的影响等问题[7]。而机器翻译与译后编辑相结合的模式既充分发挥机器翻译的速度, 也充分发挥人工翻译的精度[8], 译者可通过译后编辑确保译文的质量, 在处理专业术语、词义选择、句式的调整、语篇衔接和语言的自然性等方面更加精确、灵活, 更具可读性。

3.1. 术语和词汇翻译

进行技术文档的英汉翻译时, 术语翻译的准确性直接影响文档质量。机器翻译在术语和词汇翻译中常出现术语错译、无法识别日常词汇在专业领域的应用而导致语义偏差等问题, 从而导致译文的准确性和专业性有所缺失。翻译时需确保术语无歧义、简明、准确。

例 1:

原文: A big use case for LangChain is creating agents.

谷歌翻译: LangChain 的一大用例是创建代理。

译后编辑: LangChain 的一大用例就是创建智能体。

分析: 面对灵活使用的多义术语, 机器翻译往往无法根据学科进行词义选择, 从而在专业领域表现欠缺。在该文档中, 谷歌翻译无法理解在这一特定上下文中“agents”的具体含义, 导致译文出现了错误, 而术语翻译错误会导致信息的错误表达, 削弱文本的专业性。在人工智能和机器学习领域, 按照中文使用习惯和在中文大多数文献里, “agents”这一专业术语称为“智能体”。因此, 对于该技术文档的翻译, 使用“智能体”这一术语表达可以更精准地反映原文的内容, 符合技术文档翻译的准确性和专业性。

例 2:

原文: By themselves, language models can't take actions—they just output text.

谷歌翻译: 语言模型本身无法采取行动——它们只是输出文本。

译后编辑: 语言模型本身不能执行操作——它们只能输出文本。

分析: 由于词对齐不当、词语搭配等原因, 机器翻译对于短语的识别和翻译可能出现偏离。谷歌翻译将“take actions”翻译为“采取行动”, 虽然字面上来看没有错译, 但在该语境中“take actions”特指AI系统“执行具体操作”, 属于技术场景下的功能表述, 该译文未结合技术语境。为了让译文更符合技术文档的语义指向, 将“采取行动”改译为“执行操作”, 更能体现技术文档的准确、清晰。

例 3

原文: All Toolkits expose a `get_tools` method which returns a list of tools.

译文: 所有工具包都公开 `get_tools` 方法, 该方法返回工具列表。

译后编辑: 所有工具包均提供 `get_tools` 方法, 该方法返回工具列表。

分析: 由于英文词义丰富、灵活, 机器翻译难以结合语境选择恰当合适的词义。谷歌翻译将“expose”译为“公开”, 不符合技术文档中组件提供方法的常用表达, 存在语义偏差。将“公开”改为“提供”, 更符合技术文档中方法可被调用、向外部提供功能接口的核心含义, 提升译文的专业性, 语言更加自然, 符合中文的表达方式。

3.2. 句式结构

英文技术文档常使用长句、被动语态、复杂修饰成分等句式结构, 以实现表达的严谨性与逻辑性。机器翻译往往直译这些句式, 导致中文译文句式结构混乱、表达生硬, 不符合中文技术文本的简洁明了的表达习惯。译后编辑需通过拆分长句、转换语态、重组结构等方式优化句式, 提升译文的可读性与流畅性。

例 4

原文: The name, description and JSON schema are provided as context to the LLM, allowing the LLM to determine how to use the tool appropriately.

谷歌翻译: 名称、描述和 JSON 模式作为 LLM 的上下文提供, 允许 LLM 确定如何正确使用该工具。

译后编辑: 工具的名称、描述及 JSON 模式为 LLM 提供参考语境, 使其合理判断该工具的使用方式。

分析: 谷歌翻译按照英语的句法结构翻译, 将“提供”作为谓语, 将“作为 LLM 的上下文”译为状语, 这种直译不符合汉语的句式结构, 可读性较差, 需要按照目的语的习惯合理安排结构[9]。译后编辑首先增译“名称、描述及 JSON 数据模式”的归属, 即“工具”, 避免指代模糊; 将英语的被动态“are provided”转换为中文常用的主动态, 译为“为 LLM 提供参考语境”, 来展现大模型通过参考上下文来决定调用的工具和传入的参数的含义。编辑后的语序符合中文表达逻辑, 句式简洁流畅。

例 5

原文: Chat models that have been fine-tuned for tool calling will be better at tool calling than non-fine-tuned models.

谷歌翻译: 针对工具调用进行了微调的聊天模型在工具调用方面会比未微调的模型表现更好。

译后编辑: 经工具调用专项微调的聊天模型, 其工具调用能力会更出色。

分析: 英语句式多呈现为勾连性, 汉语多呈现为块状性和离散性特点[10]。原文使用现在完成时的被动态“have been fine-tuned”, 谷歌翻译直译为“进行了微调的”, 导致译文表述繁琐, 且“在工具调用方面会比……表现更好”句式重复。改进译文将被动态转为主动语态“经……专项微调的”, 简洁明了; 同时, 译后编辑不拘泥于原文结构, 将原文中的多个介词转换为小句, 并将“在工具调用方面会比……表现更好”优化为“其工具调用能力会更出色”, 避免“微调”两个字的反复出现, 结构及语义更清晰, 符合技术文档的清晰、简洁的要求。

例 6

原文: Non fine-tuned models may not be able to use tools at all, especially if the tools are complex or require multiple tool calls.

谷歌翻译: 未微调的模型可能根本无法使用工具, 尤其是当工具很复杂或需要多次工具调用时。

译后编辑: 未经过微调的模型可能完全无法使用工具, 尤其是面对复杂工具或需要多次调用工具的场景。

分析: 原文中 “if the tools are complex or require multiple tool calls” 是条件状语从句, 谷歌翻译将其直译为 “当工具很复杂或需要多次工具调用时”, 表达生硬, 有 “翻译腔”, 且 “需要多次工具调用”的逻辑主语不清晰。因此, 译后编辑将条件状语从句重组为 “面对复杂工具或需要多次调用工具的场景”, 增译 “面对……的场景” 明确了逻辑关系, 句式更符合中文表达习惯; 同时 “完全无法” 比 “根本无法” 更贴合技术文档的正式语气, 符合整体行文风格的一致性。

3.3. 语篇衔接

语篇衔接与连贯程度会影响文档信息传递的流畅性与可读性。机器翻译往往忽视语篇层面的逻辑关系, 导致指代模糊、句式呼应缺失、逻辑连接词使用不当等问题。译后编辑需通过明确指代、强化句式呼应、优化逻辑连接词等方式, 增强语篇的衔接。

例 7

原文: Agents are systems that use an LLM as a reasoning engine to determine which actions to take and what the inputs to those actions should be. The results of those actions can then be fed back into the agent and it determines whether more actions are needed, or whether it is okay to finish.

谷歌翻译: 代理是使用 LLM 作为推理引擎来确定要采取哪些行动以及这些行动的输入应该是什么的系统。然后可以将这些行动的结果反馈给代理, 并确定是否需要更多行动, 或者是否可以完成。

译后编辑: 这类智能体是将 LLM 作为推理引擎, 进而确定需采取的行动及相应输入的系统。随后, 这些行动的结果可反馈给该智能体, 由其判断是否需要进一步行动, 或是已可结束流程。

分析: 谷歌翻译中的最后一句 “并确定是否需要更多行动, 或者是否可以完成” 存在主语省略错误。原文中 “it determines” 的主语是前文明确的 “agent”, 但译文省略了主语, 语篇的衔接出现问题, 造成语义模糊。译后编辑则通过 “其” 明确主语是前文的智能体的指代, 逻辑关系明显, 语义衔接自然流畅。另外, 在结尾增译 “流程” 二字, 使完成的对象更具体, 强化呼应前文的动作逻辑, 上下文更加连贯。

例 8

原文: If you are still using AgentExecutor, do not fear: we still have a guide on how to use AgentExecutor. It is recommended, however, that you start to transition to LangGraph. In order to assist in this we have put together a transition guide on how to do so.

谷歌翻译: 如果您仍在使用 AgentExecutor, 请不要担心: 我们仍然有关于如何使用 AgentExecutor 的指南。但是, 建议您开始过渡到 LangGraph。为了协助您, 我们整理了一份关于如何过渡的指南。

译后编辑: 如果您仍在使用 AgentExecutor, 请不要担心: 我们仍然有关于如何使用 AgentExecutor 的指南。但是, 建议您开始过渡到 LangGraph。为协助您完成这一过渡, 我们整理了对应的过渡指南。

分析: 原文 “this” 明确指代前文的 “transition to LangGraph”, 意思是协助用户完成过渡, 而谷歌翻译的 “为了协助您” 未点明协助的具体事项, 导致这与前文建议过渡的语义衔接逻辑不清晰。因此译后编辑明确该代词的含义, 补充 “完成这一过渡” 的表达, 明确协助的具体内容, 达到句式前后呼应的效果, 并在后文用 “对应的” 替代重复表述, 使语篇表达更自然, 逻辑衔接更清晰。

例 9

原文: AgentExecutor was essentially a runtime for agents. It was a great place to get started, however, it was not flexible

enough as you started to have more customized agents.

谷歌翻译: AgentExecutor 本质上是代理的运行时。它是一个很好的起点, 但是, 当您开始拥有更多定制代理时, 它不够灵活。

译后编辑: AgentExecutor 本质上是代理的运行时。它曾是入门的理想选择, 但当您需要使用更多定制化代理时, 其灵活性就显得不足了。

分析: 首先, 在谷歌翻译中, “它是一个很好的起点”表述笼统, 忽略了原文的时态而导致和后文的逻辑衔接不流畅, 因此译后编辑强调其“曾是入门的理想选择”, 和后文的不够灵活形成连贯性逻辑。其次, 谷歌翻译中的“它不够灵活”和前文缺乏条件和结果的衔接词, 导致句式松散, 逻辑关联不紧密。因此, 译后编辑新增“就”字, 形成“但当……时, 其灵活性就显得不足了”的句式, 使前后分句的衔接更紧凑。

4. 总结

本文本研究聚焦英汉翻译方向的技术文档, 以 LangChain 官方技术文档及谷歌翻译结果为研究对象, 探究了技术文档的机器翻译及译后编辑策略。研究发现机器翻译在技术文档翻译中虽能提升效率, 但在术语翻译、句式结构、语篇衔接等三个维度存在不足。因此需将机器翻译和译后编辑相结合, 利用机器翻译的速度和人工翻译的深度, 保证技术文档的翻译质量, 为技术文档翻译提供了高效且经济的解决方案。本研究结论仅适用于软件开发类文档的汉译, 未来研究可从三方面展开, 一是拓展文本范围, 二是对比不同类型软件机器翻译译文质量, 三是通过量化研究构建译文质量评估体系。在技术文档翻译的重要性越来越突显的时代, 高质量的机器翻译为技术知识的跨语言传播提供有力支撑, 同时也为构建无障碍交流的信息社会提供保障。

基金项目

本研究得到武汉工程大学第十六届研究生教育创新基金(CX2024197)、研究生教育教学重点建设项目(2022ZDXM10)以及 2025 年上海外教社信息技术有限公司“教育部产学合作协同育人项目”资助。

参考文献

- [1] 任文. 机器翻译伦理的挑战与导向[J]. 上海翻译, 2019(5): 46-52+95.
- [2] 杨文地, 范梓锐. 科技语篇机器翻译的译后编辑例析[J]. 上海翻译, 2021(6): 54-59.
- [3] 李双燕, 孙晔芃, 李欣春. 技术写作与翻译的融合: 现状调查与教学启示[J]. 外语界, 2022(2): 63-71.
- [4] 李双燕. 国内技术文档翻译效果量化研究[J]. 外国语文, 2017, 33(1): 132-139.
- [5] 嵇大伟. 技术文档编辑之当下论域与本土革新研究[J]. 出版科学, 2014, 22(2): 82-85.
- [6] 沈梦菲, 黄伟. 寻找机器翻译痕迹——神经机器翻译文本的句法特征研究[J]. 外语教学与研究, 2024, 56(3): 429-441+480-481.
- [7] 王珍珍. 数字时代机器翻译引发的问题及思考[J]. 云南师范大学学报(哲学社会科学版), 2024, 56(5): 44-54.
- [8] 崔启亮. 论机器翻译的译后编辑[J]. 中国翻译, 2014, 35(6): 68-73.
- [9] 蔡力坚. 科技类文本的翻译探析[J]. 中国翻译, 2023, 44(6): 154-157.
- [10] 赵朝永, 王文斌. 汉语流水句英译的结构转换策略: 英汉时空性强弱差异视角[J]. 外语教学, 2023, 44(4): 15-22.