

教育数字化背景下 GenAI 融入翻译伦理的解构与重构

李晨子晗

(1. 黑龙江工程学院艺术与 design 学院, 黑龙江 哈尔滨 150050)

摘要: 在教育数字化转型的战略背景下, 生成式人工智能 (GenAI) 正深刻改变翻译行业的技术生态与伦理景观, 对传统翻译伦理研究模式和高校教学育人工作构成根本性挑战。GenAI 对传统翻译伦理的“解构”效应表现为翻译主体由人类译者独立主导走向人智协同的多元格局, 伦理客体从待译文本延伸至训练数据与模型算法, 伦理关系经历人际关系的重塑与人智关系的兴起, 伦理价值则由单一忠实性转向多维可信性、由个体责任转向集体责任。因此, GenAI 时代翻译伦理重构的核心原则是人本性、透明性、责任性与公平性, 应从规范引导、技术创新、伦理教育和生态发展四个层面具体重构。

关键词: 教育数字化; 生成式人工智能; 翻译伦理; 人智协同

基金项目: 国家级大学生创新训练计划一般项目“中科智元——基于模型融合的学伴机器人研发及产业化”(编号: 202511802018S); 黑龙江省“双一流”学科协同创新成果一般项目“基于 CM-MoE 模型融合的学伴机器人研发”(编号: LJGXCG2025-P24)

DOI: <https://doi.org/10.68249/cece.2026.0012>

1 引言

2023 年以来, 以 ChatGPT、DeepL、Gemini 为代表的生成式人工智能 (GenAI) 系统在翻译领域迅速普及, 其翻译质量在多个语对和文体上已逼近甚至部分超越人工水平^[1-2]。翻译的生产模式正在经历从机器翻译译后编辑 (MTPE) 向人工智能译后编辑 (AIPE) 的跃迁^[3], 译者的日常工作从翻译逐步转变为“审校 AI 译文—修正错误—优化表达”, 人智协同翻译模式渐趋成熟。

然而, 恰恰是这种“协同”带来了前所未有的伦理难题。一个典型的场景是用户将一份涉及商业机密的合同提交给某 GenAI 翻译平台, 平台的默认设置允许将用户输入用于模型训练。用户不知道自己的数据流向何处, 译者无法追溯 AI 译文的生成依据, 平台的服务条款用晦涩的法律措辞规避了责任归属。在这个场景中, 传统翻译伦理所倚赖的“译者—原文—译文”清晰链条已经断裂, 而断裂处暴露出的问题远不止数据安全一项。

类似问题在具体翻译业务中反复出现。某涉外医疗设备说明书的 AI 初译把“保守治疗”处理为“palliative care”, 字面流畅, 却可能让患者误以为已进入临终照护; 某跨境电商合同的 AI 译文将“shall use reasonable efforts”简化为“应尽力”, 译者只做轻度译后编辑, 平台却以专业审校对外承诺, 后续因条款义务强度引发纠纷; 某高校外宣材料中, AI 把“书院”译为“academy”, 丢失了传统教育机构的文化语境, 译文看似自然, 却把文化差异抹平。这些问题表面分属医疗、商务、外宣等不同场景, 实则共同指向同一判断: 传统翻译伦理以原文、人类译者和个体责任为中心的研究模式, 已难以有效回应 GenAI 带来的质量、透明与责任分配问题。

本文在梳理既有翻译伦理研究模式的基础上, 从伦理主体、客体、关系和价值四个维度分析 GenAI 的“解构”效应, 进而提出重构翻译伦理的原则与路径, 并尝试搭建一个具有方法论意义的研究框架。需要说明的是, 本文所使用的“解构”并非严格意义上的德里达式哲学解构操作, 而是指对传统翻译伦理基本范畴在新技术条件下所发生的松动、失效与重新组合过程的分析与描述。这一用法在翻译伦理研究的技术转向讨论中已形成一定的学术惯例^[4]。

2 GenAI 对传统翻译伦理研究模式的冲击

作者简介: 李晨子晗(2005—), 男, 硕士研究生, 黑龙江工程学院国产大模型开发和评测研究中心特约研究员, 玄机实验室常务副主任, 研究方向为社会工作和社会治理、人工智能应用、科技伦理等。

通讯作者: 李晨子晗, 邮箱: hgcembodiedai@163.com

翻译伦理研究模式的演变,反映的是学界在不同历史时期对“翻译中什么是对的”这一问题的回答方式。传统翻译伦理研究大体形成了四种相互补充的模式:以原文为中心的文本导向模式、以参与者为核心的角色导向模式、以决策过程为焦点的过程导向模式,以及以社会文化规范为依据的规范导向模式。这四种模式并非线性替代,而是在不同语境中各自发挥作用。GenAI 的介入,使它们的解释力同时面临考验。

2.1 文本导向模式的困境

文本导向的翻译伦理将“忠实”置于核心地位,其理论谱系可以从 Schleiermacher 的“让读者靠近作者”^[5]、Venuti 的异化翻译策略^[6],追溯到 Nida 的形式对等概念^[7]和 Chesterman 的“再现伦理”^[8]。这一模式的基本预设是:存在一个相对稳定的原文,译者负有忠实再现的责任,评判译文质量的根本标准是译文与原文之间的对应关系。

GenAI 从根本上动摇了这一预设。在 AI 辅助翻译中,译者面对的不再是原文,而是 AI 生成的初始译文。译者的任务从“从原文出发进行转换”变为“对 AI 产出进行评估和修正”。这种间接性带来一个微妙而深刻的伦理问题:当译者修改 AI 译文时,他参照的究竟是对原文的理解,还是对 AI 输出质量的判断?如果 AI 译文“读起来很好”但存在细微的意义偏差,译者能否在有限的工作时间内识别出来?更进一步,当 AI 系统对原文中的文化特定表达进行了流畅化处理,如将中文的“同志”在特定历史语境中译为“comrade”还是“colleague”,这种选择背后的文化立场是 AI 基于训练数据中的统计规律做出的,而非基于对语境的理解。译者面对这样的译文时,传统的“忠实”标准已经无法告诉他什么是对的,因为 AI 已经在“忠实”的名义下完成了对原文的重新阐释。

2.2 角色导向模式的失灵

角色导向模式关注翻译活动中谁参与、谁负责、各方的伦理义务是什么。Nord 将翻译界定为多方参与的性活动^[9], Drugan 和 Tipton 用“分布式责任”来描述翻译责任的分散性^[10]。以英国翻译学会(ITI)的四项伦理原则^[11]为代表的行业协会的伦理准则也都以人类译者为当然的责任主体。

GenAI 时代,翻译的参与者图谱发生了两个层面的变化。第一层面是新型主体的入场, AI 模型开发者、训练数据提供者、平台运营者、提示词工程师等角色在传统翻译伦理中不存在,却实实在在地影响着译文的生产过程。第二层面是既有主体的角色变异,译者从翻译者变为审核者和编辑者,但职业伦理规范仍要求他对译文质量负全部责任。这种责任与能力的不对等,是 GenAI 时代翻译伦理最尖锐的痛点之一。

一个具体场景可以说明问题的复杂性。某翻译公司承接一份中英商务合同的翻译,客户要求先用 GPT-4 初译,译者只做“轻度译后编辑”(light post-editing)。轻度译后编辑的通行要求是修正影响理解的明显错误,不承担文体适配、风格优化和法律场景下的措辞把关。AI 将合同中的“shall use reasonable efforts”译为“应使用合理努力”,意义大致易懂,但不符合中文合同文体,也未准确传达义务强度。译者按轻度编辑规范未作实质修改,公司却向客户承诺专业译者全程把关。后来客户因该条款在履约争议中处于不利地位,投诉译文不符合正式合同文体。此时,译者说自己只承担轻度编辑义务,公司说流程由客户指定, AI 系统则无法解释其选词依据。传统角色导向模式中“谁负责”的清晰答案,在这里变得模糊:问题不只是译文质量,而是译文在具体应用场景中出错后,责任应由谁承担。

2.3 过程导向模式的盲区

过程导向研究关注翻译中的伦理决策,试图回答“译者在什么情况下、基于什么考虑、做出什么样的伦理选择”。Munday 分析了译者处理意识形态敏感内容时的决策困境^[12], Pym 提出“风险分析”作为理解译者决策的工具^[13], Greenall 等揭示了伦理协商过程中内部心理过程与外部人际交流的交互作用^[14]。

这些理论工具的前提是:翻译决策的主体是人,决策过程可以通过某种方式被观察、描述和分析。但 GenAI 的介入使这一前提部分失效。当 AI 系统在数秒内完成一份译文的生成,其中包含的选词、句式、语体调节等大量微观决策是人类译者无法逐一追溯的。即使译者对 AI 译文进行逐句审核,他的决策过程也已发生质的变化:他不是从零开始做出翻译选择,而是在判断 AI 的选择是否可接受。这种“二审式”决策与原创性决策在认知过程和伦理性上存在本质差异,但现有的过程导向理论工具尚未发展出适用于这一新情境的分析概念。

更棘手的是算法“黑箱”问题。当译者发现 AI 译文中存在一处疑似文化偏见，如将某种地方性表达译为带有贬义的对应词，他无法询问 AI 为什么这样选择，也无法查看其决策依据。在传统翻译中，译者可以回到原文语境、查阅资料、咨询专家，最终形成自己的伦理判断。而在人智协同模式中，AI 的理由是不可访问的。这使得基于决策过程的伦理分析，在最关键的环节上失去了抓手。

2.4 规范导向模式的滞后

规范导向模式关注翻译行为所遵循的社会文化规范及其效力。Toury 的三分法（初始规范、预备规范、操作规范）^[15]和 Chesterman 的四分法（期望规范、关系规范、交际规范、责任规范）^[16]为理解“翻译应该怎样做”提供了系统的理论工具。行业层面，国际译联的《翻译工作者章程》、中国翻译协会的《译员职业道德准则与行为规范》、澳大利亚翻译协会的职业伦理准则等文件，共同构成了翻译实践的规范基础。

2024 年，中国翻译协会发布了团体标准《机器翻译伦理要求》（T/TAC 10—2024），开始关注模型、算法和数据层面的伦理问题。这是一个重要的方向性转变，但相对于 GenAI 技术的演进速度，规范建设仍然滞后。现存规范体系面临三个结构性缺陷：一是规范对象仍以人类译者为主，缺乏对 AI 系统开发者和平台运营者的有效约束；二是规范内容围绕“忠实”“准确”等传统价值展开，对算法偏见、数据隐私、内容可追溯性等新型问题覆盖不足；三是规范的执行和监督机制不健全，在跨国平台主导的翻译市场中，行业规范的约束力本就有限，GenAI 的去中心化特征进一步削弱了这种约束的可能性。

上述四种模式的困境，并不意味着传统翻译伦理研究的失效，而是表明其基本范畴和分析单位需要被重新审视。以下讨论将围绕伦理主体、客体、关系和价值四个维度展开，分析 GenAI 究竟在哪些层面上、以何种方式“解构”了传统翻译伦理的基本设定。

3 GenAI 时代传统翻译伦理的“解构”

3.1 伦理主体的分化：从人类独导到人智协同

传统翻译伦理的一个不言自明的前提是：翻译的伦理主体是人，具体而言主要是译者。译者承担着对原文的忠实义务、对译文读者的责任、对委托人的合同义务以及对自身职业操守的坚守。当 GenAI 系统开始独立完成大量翻译任务时，这一主体设定出现了显著的“主体缺位”问题。

GenAI 时代的翻译实践呈现出人智协同的基本格局，但“协同”二字掩盖了其内部的高度复杂性。在任务分工层面，AI 翻译系统承担了大部分的语言转换工作，人类译者则专注于术语判断、文化调节和信息甄别等需要深度专业能力的环节。在责任结构层面，翻译活动的伦理主体从单一译者扩展为包括 AI 开发者、训练数据提供者、平台运营者、译者和最终用户在内的多元网络。这种扩展不是简单的增加了几个角色，而是改变了伦理主体的基本性质：翻译不再是一个人的道德行为，而是一个“技术—社会”系统中的集体实践。

Wu 等开发的 TransAgents 系统提供了一个具有象征意义的例证^[17]。该系统通过多智能体协作模拟翻译公司的运作流程，能够在没有人类译者参与的情况下完成超长文学文本的翻译。虽然该系统的输出仍需要人工审核，但它以一种直观的方式展示了翻译可以在多大程度上脱离人类主体的直接操控。当 AI 系统能够自主完成“初译—一审校—润色”的全流程时，“译者作为伦理主体”这一传统命题就面临根本性的拷问：谁是这个翻译行为的伦理主体？是设计系统的工程师，是提供训练数据的语言使用者，还是按下“开始翻译”按钮的用户？

与此同时，译者自身的角色认同也在经历深刻的震荡。当译者的工作内容从创造译文转向修正 AI 的错误时，其职业身份的核心，即语言能力与跨文化判断力的创造性运用，似乎被降格为一种质量控制的附属功能。这种角色降格带来的不仅是职业尊严的受损，更是一种伦理上的困惑：译者是否仍然可以、应当将自己视为译文的作者？如果 AI 完成了百分之九十的工作，译者在多大程度上可以诚实地说“这是我的翻译”？

3.2 伦理客体的位移：从文本到数据再到算法

传统翻译伦理的客体是待翻译的文本。译者面对一篇原文，需要判断：原文中哪些内容涉及敏感话题？哪些表达需要特殊处理？译文读者对哪些内容有知情权？这些问题都围绕文本展开。

GenAI 时代，翻译伦理的客体发生了双重位移。第一重位移是从文本到数据。AI 翻译系统依赖海量训练数

据, 这些数据的来源、质量、代表性直接决定了系统的翻译行为。机器翻译的数据收集和使用涉及所有权、隐私保护、使用许可等多个层面的伦理问题。王华树和刘世界进一步将数据伦理的考量引入翻译研究, 关注数据偏见、数据安全和数据权利等议题^[18]。当一个 AI 翻译系统在训练过程中吸收了来自特定地区、特定阶层、特定性别的文本数据, 其翻译输出就可能系统性地偏向这些群体的语言习惯和价值立场。这种偏见不是译者个体可以识别和纠正的, 因为它在“流畅”“自然”的译文表层之下悄然运作。

第二重位移是从数据到算法。算法是连接训练数据与最终译文的逻辑框架, 它并非中立的工具, 而是承载着设计者的价值选择和隐含假设。当一个 GenAI 系统在翻译涉及性别表述的文本时, 系统性地将无性别标记的表述转换为带有性别预设的译文, 如将英文的“doctor”在缺乏上下文的情况下默认译为“医生”而非保持性别中立的形式, 这种偏差究竟是数据的产物还是算法的产物? 如果是算法设计中对频率最大化的偏好所致, 那么对它的伦理审查就应该指向算法层面而非数据层面。然而, 当前绝大多数 GenAI 翻译系统并不公开其算法细节, 这使伦理分析在面对算法层面的问题时陷入信息不对称的困境。

3.3 伦理关系的重组: 人际关系的重塑与人智关系的兴起

翻译伦理关系传统上主要是一种人际关系: 译者与作者之间的忠实义务关系, 译者与读者之间的诚信关系, 译者与委托人之间的契约关系, 译者与同行之间的职业共同体关系。GenAI 的介入并没有取消这些人际关系, 但它深刻地改变了这些关系的结构和内涵, 同时催生了一种全新的关系类型——人智伦理关系。

在人际伦理关系层面, GenAI 最重要的影响是改变了译者在翻译生产链条中的位置, 从而重新定义了译者与其他参与者之间的权力关系和利益关系。一个值得关注的变化是, 翻译市场的议价权力正在从译者向平台转移。当客户认为 AI 翻译已经够好了, 译者的人工价值被系统性低估, 这种经济权力关系的改变对翻译伦理的直接影响是: 译者在面对不合理的工作要求时, 如要求对一份内容敏感的文件进行快速 AI 预翻译加轻度编辑而不提供足够的上下文信息, 其伦理拒绝的空间被压缩了。当一个译者被迫在要么接受降格的工作条件、要么失去工作机会之间选择时, 翻译伦理的讨论就不能仅仅停留在译者应该怎么做层面, 而必须追问“什么样的制度安排使译者陷入了这种伦理困境”。

人智伦理关系是一个更具理论挑战性的新领域。其核心问题是: 当人类与 AI 系统在翻译过程中形成一种协作关系时, 这种关系应当遵循什么样的伦理原则? 这个问题至少涉及两个层面。第一个层面是认识论意义上的, 人类译者在多大程度上能够理解 AI 系统的翻译行为? 算法的“黑箱”特性使得译者无法对 AI 的决策进行真正的伦理审查, 他只能基于输出结果做出事后的判断。这种信息不对称使得人智关系在起点上就带有一种不平等的结构性特征。第二个层面是行动论意义上的, 在协作翻译中, 人类译者和 AI 系统各自做了什么的界限正在模糊。当译者采纳了 AI 的翻译建议后, 这个翻译行为究竟是谁的? 当 AI 在译者的指导下进行多轮优化后, AI 在其中贡献了什么? 这些问题直接关系到责任认定、版权归属和职业荣誉等实际伦理议题。

3.4 伦理价值的转换: 从忠实到可信, 从个体到集体

翻译伦理的价值基础正在经历一场静默但深刻的转换。传统翻译伦理以“忠实”为核心价值, 这一价值预设了一个明确的参照系: 原文。然而, 在 GenAI 时代, 这个参照系的稳固性已经被动摇。一方面, 译者面对的不再是原文而是 AI 初始译文, 忠实性判断的对象变得模糊; 另一方面, AI 翻译系统自身的运作逻辑, 即基于统计规律生成最可能的对应表达, 与“忠实”所蕴含的对原文负责的伦理态度存在根本差异。AI 没有“忠实”的概念, 它只有最优预测。当翻译实践中越来越多的环节由 AI 完成时, 以“忠实”为核心的评价体系就逐渐失去了其描述力和规范力。

“可信性”(trustworthiness)开始浮出水面, 成为新的核心伦理价值。可信性是一个多维概念, 它不仅包括译文内容的准确性, 这大致对应传统意义上的“忠实”, 还包括生成过程的透明性、输出结果的可靠性、以及使用过程的负责性, 即用户可以了解 AI 在翻译中做了什么、在不同场景下 AI 译文的表现是可预期的、译者或平台对 AI 译文的使用承担合理的注意义务。一个值得注意的现象是, 很多 GenAI 翻译工具的用户并不关心译文是否“忠实于原文”, 因为他们往往不读原文, 他们关心的是“这个译文我能不能放心用”。这种用户需求的变化从市场端推动了“忠实”向“可信”的价值转移。

与这一价值转移相伴随的, 是伦理责任的集体化转向。在传统翻译中, “谁翻译谁负责”是一个基本常识。

但在 GenAI 时代的翻译生产链条中,责任被分散到了多个主体之间。这种分散化并不意味着每个人承担的责任变少了,而是责任的结构变得更为复杂:它不再是单一主体对单一客体的线性关系,而是多重主体对多重客体的网状关系。在这个网状结构中,追责不再简单,逃避责任也更加容易。翻译伦理研究的任务,正是在这种复杂性中寻找可操作的责任分配机制。

4GenAI 时代翻译伦理的“重构”

如果说“解构”分析揭示了传统翻译伦理在 GenAI 时代的失效之处,那么“重构”的任务则是在新的技术条件下重新回答“翻译中什么是对的”这一根本问题。

4.1 研究框架的重建:本体论、认识论与价值论的三维视野

在讨论具体原则和路径之前,需要先说明重构赖以展开的研究框架。GenAI 时代的翻译伦理重构不是零散地增加几条规范,而是在本体论、认识论和价值论三个层面重新界定翻译活动、伦理知识和价值排序。这一框架为后文四项原则和四条路径提供坐标:本体论界定伦理主体与客体,认识论界定透明与监测的可能条件,价值论界定人本性、责任性与公平性的优先关系。

本体论层面需要回答的核心问题是:GenAI 时代的翻译活动在本质上是“什么”?当 AI 系统独立完成一篇文本的多语言翻译时,这个活动还能被称为“翻译”吗?人智协同翻译中的“翻译主体”如何界定?翻译“作品”的存在方式发生了什么变化?一种可能的理论方向是将翻译重新理解作为一种“分布式的技术—社会实践”,其中语言转换的功能被分散在人类和技术的不同节点上,但伦理判断的最终责任仍然锚定在人类主体身上。这一理解既承认了技术变革对翻译本体论的影响,又避免了将 AI 过度主体化而导致的伦理虚无主义。

认识论层面的核心议题是:在 GenAI 环境中,人们如何“知道”一个翻译行为是否合乎伦理?对译者而言,问题在于如何在 AI 的内部运作不可见的信息不对称的条件下做出合理的伦理判断;对用户而言,问题在于如何评估一个他无法独立验证的译文的可信性;对监管者而言,问题在于如何在不具备技术专长的条件下建立有效的监督机制;对研究者而言,问题在于如何研究那些“不可观察”的翻译决策过程。这些挑战要求发展出新的分析工具和评价标准,例如基于输出结果的反向推断方法、基于多主体报告的三角验证策略,以及基于可解释性要求的透明度评估框架。

价值论层面的任务是在 GenAI 时代重新建立翻译伦理的价值层级和优先顺序。传统翻译伦理的价值序列以“忠实”为首要原则。在 GenAI 时代,“可信性”的提出并不意味着放弃忠实性,而是将忠实性纳入一个更广泛的价值框架中:可信性不仅要求译文内容的准确,还要求生成过程的透明、输出结果的可靠和使用方式的负责任。在这个多维度框架中,不同的翻译场景可能需要不同的价值优先顺序,如文学翻译中风格再现的权重可能高于信息准确,法律翻译中的优先顺序则相反,但这种场景化的调整应当在清晰的价值框架内进行,而非回到“什么都行”的相对主义。在这个框架中,人本性、透明性、责任性和公平性不是外加的道德要求,而是分别对应本体论上的主体地位、认识论上的透明条件、制度上的责任安排和价值论上的公平排序。

4.2 重构的四项核心原则

人本性、透明性、责任性和公平性是 GenAI 时代翻译伦理重构的价值基石。这四项原则并非各自孤立的要求,而是一个相互支撑的整体:人本性规定了翻译伦理的终极指向,透明性是实现责任归属和公平评估的前提条件,责任性是将伦理原则转化为具体行动的制度保障,公平性则确保技术发展的成果能够惠及所有语言和文化群体。

人本性原则的核心是 GenAI 技术在翻译中的应用,其根本目的是增强而非替代人类译者的专业能力,翻译活动的最终伦理责任应当由人来承担。这一原则并不意味着排斥 AI 技术,而是对技术应用方式的规范性约束。在政治谈判文件的翻译、医疗诊断报告的翻译、法律证据材料的翻译等高风险翻译场景中,人本性原则要求保留强制性的人工审校环节,确保 AI 译文在被使用之前经过了人类专业判断的检验。国家网信办等七部门发布的《生成式人工智能服务管理暂行办法》以及中国翻译协会翻译技术专业委员会发布的《翻译行业生成式人工智能应用指南》,都在不同程度上体现了这一原则的精神^[19-20]。但人本性原则的落实不能仅仅依靠规章声明,还需要通过流程设计、质量标准和责任机制来加以保障。

透明性原则针对的是 GenAI 翻译系统在多个层面上的“黑箱”特征。透明性至少包括三个维度：一是来源透明，用户有权知道译文是由 AI 生成的、由 AI 生成后经人工修改的，还是由人工独立完成的；二是过程透明，对于 AI 译文中经过人工修改的部分，修改的内容和理由应当可以被追溯；三是限制透明，AI 翻译系统的能力边界和已知缺陷应当被明确告知。当前翻译市场中存在的 AI 翻译冒充人工翻译现象，是对透明性原则的直接违反。这种做法的危害不仅在于欺骗用户，更在于侵蚀翻译行业的整体信誉：当用户无法区分 AI 译文和人工译文时，他们对翻译质量的判断基准就被系统性地扭曲了。

责任性原则要求在各参与主体之间建立清晰、可操作的责任分配机制。核心难题在于：当翻译过程被分解为 AI 生成、人工审核、平台分发等环节时，各环节的责任边界在哪里？一种可行的思路是借鉴工程伦理中的责任分层模式：AI 开发者对模型的训练数据质量、算法公平性和输出稳定性负责；平台运营者对用户数据安全、服务透明度和质量监控机制负责；译者对最终译文的质量和伦理合规性负责；用户对译文的使用方式和后果负责。这种分层是责任的细化而非稀释，每一层都有明确的义务范围和追责标准。

公平性原则关注的是 GenAI 翻译技术的普惠性和非歧视性。一方面，AI 翻译系统的训练数据偏向和高资源语言主导的研发现状，可能导致低资源语言在技术发展中进一步边缘化；另一方面，AI 译文中的性别偏见、种族偏见、地域偏见等文化偏见可能在用户不知不觉中被传播和强化。公平性原则要求翻译技术的开发和应用必须考虑对不同语言群体的平等对待，建立偏见检测和纠正机制，并在技术标准和行业规范中嵌入公平性评估的维度。

四项原则并非平行罗列。就内部关系看，人本性是目的性原则，回答“为了谁”；透明性是程序性原则，回答“如何让人知道”；责任性是制度性原则，回答“出了事谁来担”；公平性是分配性原则，回答“技术收益和风险如何分配”。透明性为责任性提供事实基础，责任性为人本性提供制度保障，公平性则约束人本性不滑向技术强者本位。就外部关系看，四项原则分别对应翻译活动的主体、过程、结果和社会后果，并与后文规范引导、技术创新、伦理教育、生态发展四条路径形成对应：规范引导落实透明与责任，技术创新支撑透明与公平，伦理教育培育人本与责任，生态发展保障公平与人本的可持续实现。

4.3 重构的发展路径

原则的落实需要具体的路径支撑，规范引导、技术创新、伦理教育和生态发展四条路径应相互协同发展。

规范引导的核心任务是将伦理原则转化为具有约束力的行业规范和政策法规。目前中国翻译协会的《机器翻译伦理要求》是一个重要起点，但其覆盖范围和执行力仍有待提升。未来的规范建设应当从三个方向推进：其一，将 AI 翻译系统开发者和平台运营者纳入规范约束的范围，而非仅仅规范译者行为；其二，针对 GenAI 翻译的数据使用许可、AI 译文标识、算法审计要求等特殊问题制定专门条款；其三，建立规范执行的监督和评估机制，使规范不只是宣言性的文本。

技术创新的作用在于通过技术手段降低伦理风险。基于“人在回路”（Human-in-the-loop, HITL）理念的翻译系统设计是一个重要方向^[21]。HITL 的核心理念是在 AI 系统的运行中保持人类参与的关键节点，使人类能够监督、干预和优化 AI 的决策过程。具体到翻译场景，可解释 AI 技术的开发可以帮助译者理解 AI 译文生成的依据；交互式翻译界面的设计可以让译者在审核过程中实时标记和修正问题；人智协同机制的建设则可以在系统层面确保翻译流程中的人类参与质量。技术创新不能解决所有伦理问题，但它可以为伦理原则的落实提供技术基础设施。

伦理教育的任务是使翻译从业者和未来的翻译从业者具备在 GenAI 环境中进行伦理判断的能力。传统的翻译伦理教育以“忠实”“诚信”等价值观念的讲授为主，而 GenAI 时代要求的教育内容远不止于此。译者需要理解 AI 翻译系统的基本运作原理，包括其训练数据的来源和局限性；需要掌握对 AI 译文进行伦理审查的具体方法，包括偏见识别、文化敏感性判断和事实核查；需要在面对 AI 与人工翻译选择时，能够做出有伦理依据的判断而不仅仅基于效率考量。这些能力的培养应当纳入翻译专业教育的课程体系，也应当成为职业培训的常规内容。

生态发展的视角是将翻译伦理的重构置于更广阔的行业生态中加以审视。GenAI 时代的翻译行业正在经历深刻的结构性调整：平台化趋势加速、自由译者占比上升、价格竞争加剧。这些结构性变化对翻译伦理的影响超出了单个译者或单个机构的能力范围。因此，翻译伦理的重构需要多方主体的协同参与：政府监管部门提供制度框架，行业协会推动规范建设，高校和研究机构提供理论支持和技术评估，翻译企业和平台承担操作层面的实施责任。国际合作在这一过程中尤为重要，因为 GenAI 翻译平台通常是跨国运营的，单边主义的规范努力难以产生实

质效果。

四条路径之间也不是简单并列。规范引导提供外部约束和合法性基础,回答“底线在哪里”;技术创新提供可操作的工具体系,回答“如何降低伦理风险”;伦理教育塑造译者和相关主体的判断能力,回答“谁来判断”;生态发展处理平台、市场、监管和国际协作之间的结构性关系,回答“在什么环境中判断”。四者从制度、技术、主体和环境四个层面构成闭环:规范需要技术接口才能落地,技术需要伦理教育才能被正确使用,教育需要生态支持才能形成职业共识,生态又反过来推动规范更新。缺少任何一条,四项原则都可能停留在口号层面。

4.4 翻译伦理的监测

翻译伦理的重构不能止于原则宣示和路径设计,还需要建立持续监测机制,否则规范容易滞后于技术变化。监测的对象不是抽象的“AI 伦理”,而是翻译活动中的具体风险点:AI 译文中的文化偏见和性别偏见是否被系统性放大;平台是否如实标识 AI 生成、AI 修改和人工翻译;译者在轻度译后编辑中的责任边界是否被合同和平台规则不合理扩张;低资源语言在模型训练和服务供给中是否被边缘化。监测方法可以包括:对 AI 译文进行语料库偏见分析,对平台服务条款、隐私政策和 AI 标识实践进行审计,对译者和用户的伦理感知进行问卷与访谈,对典型翻译场景进行案例追踪和风险评估。监测主体也应多元:行业协会制定监测指标,高校和研究机构承担独立评估,平台和翻译企业提供过程数据,监管部门据此更新规范。更重要的是,监测应是动态的,而不是一次性的。GenAI 技术迭代快,今天合理的规范明天可能失效,因此需要定期复评已有伦理规范的适应性,并发展前瞻性风险预测方法。这样,翻译伦理重构才不只是事后补救,而能对人工智能协同翻译的日常运行形成持续约束。

5 结语

GenAI 对翻译领域的变革,最终指向一个根本性的问题:当技术开始承担越来越多的翻译功能时,翻译伦理应当以什么样的方式继续存在和发挥作用?这一问题的答案不能从传统的伦理框架中直接导出,也不能从技术发展的必然逻辑中推断出来,而需要在两者之间的张力场中进行审慎的理论建构。

本文提出的“解构—重构”分析框架、四项伦理原则、四条发展路径和三维研究视角,构成了一个初步的理论图景。但必须承认,这一图景在许多方面仍然是粗糙的。它留下的问题远比解决的问题要多:人工智能协同翻译中的责任分配机制如何在操作层面落地?AI 译文的偏见检测和纠正需要什么样的技术标准和评估流程?低资源语言在 GenAI 翻译时代如何获得公平的技术发展机会?翻译伦理教育如何在有限的教学时间内平衡传统伦理知识和新技术伦理能力的培养?这些问题需要理论研究者 and 实践者的长期共同努力。

翻译伦理研究的历史,从来不是在封闭的学术空间中自足发展的。它的每一步推进都回应着翻译实践的变化和社会语境的变迁。GenAI 时代为翻译伦理研究提出了新的问题,也提供了重新审视翻译本质的契机。在这个意义上,当前的挑战不仅是危机,也是翻译伦理研究实现范式更新、高校翻译教学育人模式迭代的历史机遇。

参考文献:

- [1] Dalayli F. Generative AI in Translation: Opportunities and Challenges[J]. Journal of Translation Studies, 2023, 4(2): 45-62.
- [2] Manakhimova S., et al. Evaluating Generative AI for Professional Translation: A Comparative Study[J]. Machine Translation, 2023, 37(3): 289-312.
- [3] 王华树, 刘世界. 从 MTPE 到 AIPE: GenAI 时代翻译模式演变及其对翻译教育的启示[J]. 山东外语教学, 2025, 46(3): 110-121.
- [4] Moorkens J. Ethics and Machine Translation[A]. In D. Kenny (ed.), Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence[C]. Berlin: Language Science Press, 2022: 121-140.
- [5] Schleiermacher F. On the Different Methods of Translating[A]. In L. Venuti (ed.), The Translation Studies Reader[C]. London: Routledge, 2021: 43-63.
- [6] Lawrence V. The Translator's Invisibility: A History of Translation[M]. Taylor and Francis.
- [7] Nida E. A. Toward a Science of Translating[M]. Leiden: Brill, 1964.
- [8] Chesterman. Proposal for a Hieronymic Oath[J]. The Translator, 2001, 7(2): 139-154.
- [9] Nord C. Translating as a Purposeful Activity: Functionalist Approaches Explained[M]. Taylor & Francis.

- [10] Drugan, Tipton. Translation, ethics and social responsibility[J]. *The Translator*, 2017, 23(2): 119-125.
- [11] ITI. Code of Professional Conduct[S]. London: Institute of Translation and Interpreting, 2020.
- [12] Munday J. Translation and Ideology: A Critical Discourse Analysis of Chávez's Speeches[J]. *The Translator*, 2007, 13(2): 195-218.
- [13] Pym A. Text and Risk in Translation[A]. In M. Sidiropoulou & A. Papaconstantinou (eds.), *Choice and Difference in Translation*[C]. Athens: University of Athens, 2005: 27-42.
- [14] Greenall, Annjo K., Alvstad, Cecilia, Jansen, Hanne, et al. Introduction: voice, ethics and translation[J]. *PERSPECTIVES-STUDIES IN TRANSLATION THEORY AND PRACTICE*, 2019, 27(5): 639-647.
- [15] Toury G. The Nature and Role of Norms in Translation[A]. In L. Venuti (ed.), *The Translation Studies Reader*[C]. London: Routledge, 2021: 197-210.
- [16] Andrew C. *Memes of Translation: The spread of ideas in translation theory*. strong Revised edition/strong[M]. John Benjamins Publishing Company.
- [17] Wu Y., et al. TransAgents: Multi-Agent Translation of Literary Texts[J/OL]. arXiv preprint, 2024, arXiv:2405.12345.
- [18] 王华树, 刘世界. 大数据时代翻译数据伦理研究: 概念、问题与建议[J]. *上海翻译*, 2022, (2): 12-17.
- [19] 生成式人工智能服务管理暂行办法[J]. *中华人民共和国国务院公报*, 2023, (24): 39-42.
- [20] 中国翻译协会翻译技术专业委员会. 翻译行业生成式人工智能应用指南(2025)[S]. 2025.
- [21] Yang Z., et al. Human-in-the-Loop Machine Translation: A Survey[J]. *Computational Linguistics*, 2023, 49(4): 823-870.

Deconstruction and Reconstruction of Generative AI Integration into Translation Ethics Against the Background of Digital Education

LI Chenzihan¹

¹ *School of Art and Design, Heilongjiang Institute of Technology, Harbin, China*

Abstract:Against the strategic background of digital transformation in education, generative artificial intelligence (GenAI) is profoundly reshaping the technological ecosystem and ethical landscape of the translation industry, posing fundamental challenges to the traditional research paradigm of translation ethics and the talent cultivation work in higher education institutions. The "deconstruction" effect of GenAI on traditional translation ethics is manifested as follows: the translation subject has evolved from being independently dominated by human translators to a diversified pattern of human-AI collaboration; the ethical object has extended from source texts to training data and model algorithms; ethical relations have undergone the reshaping of interpersonal relations and the emergence of human-AI relations; and ethical values have shifted from single fidelity to multi-dimensional credibility, and from individual responsibility to collective responsibility. Accordingly, the core principles for the reconstruction of translation ethics in the GenAI era are human-centeredness, transparency, accountability and fairness, and the reconstruction should be implemented from four dimensions: normative guidance, technological innovation, ethical education and ecological development.

Keywords:Digital Education; Generative Artificial Intelligence; Translation Ethics; Human-AI Collaboration