

特稿

人工智能辅助翻译中的人类能动性与责任归属

孙艺风

(澳门大学 人文学院, 澳门 999078)

摘要:本文探讨人工智能在笔译与口译领域的快速整合对翻译活动的本质、译者主体性及职业责任所带来的根本性挑战。文章指出,关于人工智能翻译的核心议题已从“机器能否产出流畅自然的目標语文本”转变为:在涉及法律、伦理、文化及制度性后果的翻译中,应由谁来对意义负责、对潜在风险作出判断并承担最终问责。基于此,本文强调,翻译不应被简化为单纯的语言替换过程,而应被理解作为一种蕴含伦理维度、深植于制度环境且需要专业判断的中介性实践。文章进一步剖析了人工智能在翻译中的双重角色。一方面,机器翻译与大语言模型已在术语管理、初稿生成、重复性内容处理及标准化文本生产等方面展现出显著的效率优势,并持续重塑着翻译劳动的结构。另一方面,机器生成文本的流畅性可能掩盖其潜在的语义偏差、语气失当、文化错位与制度风险,从而使得“人的痕迹”——即人类译者依据具体语境进行阐释、权衡、修正与担当责任的能力——变得愈为关键。因此文章主张,人工智能并未消除对人类译者的需求,而是改变了人类能动性的实践形态:翻译活动正日益演变成为一种人机协作系统,但人类译者必须承担其中规范性判断与责任的核心。

关键词:人工智能翻译;译者能动性;问责性;人机协作;伦理判断

中图分类号:H059 **文献标志码:**A **文章编号:**1674-6414(2026)03-0001-11

DOI:10.26922/j.cnki.waiguoyuwen.2026.03.001

0 引言

人工智能为笔译与口译领域带来的核心问题,已不局限于流利性——机器生成流畅文本的能力已标志着人机在此方面的竞争基本告一段落。更深层次的议题在于:意义能否脱离其实际效用的责任而被孤立地处理?当人工智能被整合进翻译工作流程,问题的实质便转向“问责”:由谁判定一项译文是否充分妥帖?当某个交际选择涉及法律、文化或制度性后果时,由谁评估其中风险?如果出现差错,责任又应由谁承担?当前,翻译的能动性确实已分布于人机协同的系统之中,但翻译决策所承载的规范性分量——即关乎判断、伦理与后果的那部分——无法交由算法承担。译者与口译员并非固守自身意义的“前数字时代遗

作者简介:孙艺风,男,澳门大学人文学院讲座教授,主要从事翻译理论、文学翻译、跨文化交际及世界文学研究。

引用格式:孙艺风.人工智能辅助翻译中的人类能动性与责任归属[J].外国语文,2026(3):1-11.



本刊微信公众号

存”,其重要性恰恰源于他们承担责任;这并非感伤式的表述,而是一种结构性判断。真正有待探讨的,并非人类是否仍然重要(答案无疑是肯定的),而是在那些往往被设计为“人工干预似可或缺”的系统中,如何有意识、制度化地维持职业责任与解释权威。

基于上述视角,本文尤其关注创造力、文学性及高风险文本翻译中的自动化边界。文章指出,文学、法律、医疗、哲学及政治敏感类文本不仅要求语义传递,更需对风格、歧义、文化回响、历史脉络与伦理后果作出可辩护的判断。人工智能可模拟语言模式,却无法真正承担使翻译选择获得正当性所必需的社会、文化与伦理责任。因而在这些领域,人类译者并非技术尚未完善前的权宜存在,而是翻译实践得以成立的前提条件。需要强调的是,人工智能时代的翻译研究与行业实践,其核心任务不应止步于“人类是否会被取代”的抽象争论,而应转向如何在日益深化的人机协作中,重建并维护译者的专业权威。为此,有必要在翻译教学、职业准则、工作流程与制度设计层面,明确哪些任务可安全委托机器,哪些判断必须保留于人类监督之下,并建立相应的责任分配与风险治理机制。故有必要阐明,翻译之所以依然是一个深具人文意义的命题,正因为它始终关乎意义、信任、判断与责任,而这些维度,至今仍无法被完全自动化。

1 AI 翻译争论的核心

关于人工智能将彻底取代人类译者与口译员的论点,往往建立在对翻译本质的过度简化之上——即将翻译过程简单归结为词汇和语法的直接对应转换,类似于一种“解码—再编码”式的技术操作。近年来,机器翻译与大规模语言模型的迅猛发展,在提升目标文本的流畅性、丰富词汇多样性乃至模仿特定文体风格等方面取得了显著进展,深刻改变了文本生产的工作流程(O'Brien, 2012; Kenny, 2022)。事实上,问题的核心并非在于机器能否生成语法正确、语句通顺的文本,而在于一个更为根本的层面:当所传递的意义具有高度复杂性,且直接关涉现实后果时,机器系统是否能够——以及在何种意义上能够——为“意义”的准确传递、阐释及其可能引发的实际影响承担责任?这无疑是一个意义重大、不容轻忽的问题。

此处所言的“意义重大”,特指那些超越表层语言转换的复杂交际场域:语境具有高度的动态性与不确定性,参与各方的意图往往并不明晰,而是层层嵌套、相互博弈乃至存在公开分歧;沟通过程又受到隐性权力关系的深刻渗透,制约着何种内容得以表达以及以何种方式被言说;同时,翻译行为本身可能触及切实的伦理挑战,例如在医疗、司法或冲突调解

等场景中关乎关键决策的传译,进而引发实际的法律与社会后果。在此类情境下,翻译的产出绝非中立的话语产品,而必须由一个具备认知能力、判断力与伦理意识的主体——即“人”——来阐释翻译决策的理据,为其带来的影响进行辨析,并承担相应责任(Inghilleri, 2012)。

需要指出的是,人类译者的主体能动性——包括批判性思考、情境判断、文化调适与伦理抉择——并非对前数字时代工作方式的简单怀恋,而是翻译与口译能够成为一种专业实践、制度环节以及承载伦理责任的社会行为的根本基石。正是人的介入,保障了在复杂语境中意义的恰当生成与责任的清晰归属。当前,人工智能技术快速渗透翻译实践,其发展速度已远超翻译研究、语言学、伦理学、社会学等人文学科以细致、准确的概念工具进行系统剖析与理论回应的能力。这种理论滞后使得公众围绕人机关系的讨论时常陷入两极:或忧虑技术替代,或对其能力过度乐观。许多探讨缺乏扎实的理论支撑,或仅止步于“机遇与挑战并存”的笼统表述,难以深入辨析技术究竟在何种程度、以何种方式重新塑造着翻译的本质、价值与责任归属。

2 人类判断的伦理与专业基础

译者的角色与一位借助人工智能辅助创作原创文本的作者有着本质区别。作者主要对表达自身意图负责,而译者则承担着跨语言转移“他人意义”的职责,且这一过程往往还需跨越不对称的文化、权力与知识转换。这一区分至关重要,因为人工智能系统并非中立工具。正如语言技术研究者所指出的,这些系统在其数据、建模选择与设计实践中,已嵌入了特定的假设、偏见与规范性取向(Bender et al., 2021)。若这些问题在人工智能辅助的单语写作中已构成相当的复杂性,那么在翻译活动中,其后果将更为严峻。因为翻译的任务不仅仅是生成“表面合理”的文本,更是在利益、期待与风险并不一致的多元主体之间进行中介。在功能主义框架下,诺德(Nord, 1997: 125-126)提出译者应遵循“忠诚原则”,即同时对原文生产者、目标读者及委托方承担伦理责任。皮姆(Pym, 2012)则从另一相容的视角,将翻译描述为一种跨文化风险管理,强调责任可在不同参与者之间分布,却不可因此消解。总需有人为翻译选择所带来的交际后果负责。在法律、医疗、外交等制度性场景中,这种问责性早已超越单纯的语言转换,延伸至专业、组织与监管层面。一旦认识到这一点,便可明确:人工智能中介的翻译是否“合格”,绝不能仅凭语言流畅度判定。更紧迫的核心问题是:当多方利益交织、得失攸关时,究竟应由谁来为意义的传递与后果承担最终责任?

随着机器生成文本的说服力日益增强,这一问题也显得愈发迫切。输出越自然流畅,语义、语用与语境上的偏移就越容易在不易察觉的情况下悄然渗透。在笔译与口译实践中,真正的风险往往并不在于明显的词汇错误,而在于隐含意义、语域、语气或语用力量上的细微改变。一个被弱化的法律术语、医疗问诊中被错置的预设,或是一场外交交流中语调的微妙变化,都可能在语句表面通顺的情况下,实质性地影响甚至改变最终结果。大型语言模型所生成的是统计意义上高概率的文本序列,而非根植于现实理解的意义,这一区别正是关于其解释能力与可靠性的争论焦点所在(Bender et al., 2020)。问题不只在在于偶尔的误译。语言技术还会从其训练条件中继承社会与政治层面的预设,这些预设会影响输出,并且恰恰因为语言显得连贯、精炼且自信,而更难以被察觉(Bender et al., 2021: 613-617)。在翻译语境下,这类偏移不仅是一种技术性瑕疵,更可能构成交际行为中的潜在扭曲,从而带来法律、伦理、文化或制度层面的后果。

正因如此,人类判断之所以不可或缺,恰恰在某种程度上正是因为人工智能已表现出相当的语言流利度本身。翻译研究早已扬弃翻译恰当性可由抽象规则决定的观念。相反,一项翻译只有在与其交际目的、目标读者以及具体情境相结合时,方能被判定为恰当(Nord, 1997)。皮姆(Pym, 2012: 135-138)从风险管理角度对翻译的阐释进一步强化了这一观点:翻译选择必须被置于文本进入现实世界后可能引发的实际后果中加以评估。机器翻译在流利度上的提升,并未消解翻译对特定主体的需求——该主体必须能够辨识隐含意义、语气、语域与文化共鸣的细微差异,并判断这些差异在具体语境下是否可接受。始终需要有人作出判断:一段表面流畅的译文是否真正契合当前目标,是否维护了相关各方的权益,以及当其面临质疑时,是否具备合理的辩护依据。依此理解,人的触感并非某种模糊的情感诉求,而是一种在不确定性中进行情境化判断,并为其结果承担责任的能力。

3 AI时代翻译能动性的变化

人工智能所改变的并非人类能动性的必要性,而是其配置方式。随着机器翻译与生成式模型日益深入地嵌入翻译工作流程,对人类译者与口译员的专业要求并未被消除,而是得以重新分配。奥洛汉(Olohan, 2011: 346-349)指出,翻译技术改变的是译者能动性得以实现的条件,而非简单地将其移除;这要求从业者能够判断何时适用自动化、自动化可被信任的程度,以及何时必须加以抵制。同样,切斯特曼(Chesterman, 2001)也强调,翻译始终是一种由义务建构起来的伦理实践,而这些义务无法转移给那些自身不能承担责任的系

统。即使在混合式工作流程中,仍然是人类从业者决定哪些任务可被负责任地委托,评估机器输出是否适用于既定目标,并对最终产出的交际效果负责。自动化并未消灭人类能动性,而是使其位置、作用范围与规范基础变得更加明晰。

我们正是基于这一视角来审视“人类是否仍然具有相关性”的问题。与其将“相关性”视为译者与口译员自身固有的属性,不如将其理解为在不断演变的社会—技术格局中被关系性建构的一种成就。那种要么在原则上抵制技术,要么将技术视为专业能力直接替代品的常见对立,实际上遮蔽了真正的关键所在。人工智能并不仅仅是为一个原本稳定的职业增添一种新工具,而是重新组织了翻译行为获得授权、受到评价以及被纳入问责的条件。故此,核心问题并非人类从业者是否能在文本生产的每个环节都超越机器,而在于这一领域能否识别、维护并更新那些在自动化日益强大的背景下依然不可或缺的判断形式、责任形式与伦理约束形式。

基于上述理解,相关性与伦理能动性的维系,需要在“人—机系统”中有意识地重新界定专业管辖权。能动性或许分布于人、界面、制度流程、提示词与模型输出之间,但责任的规范性锚点依然在于人类。译者与口译员的独特价值,不仅在于其生成目标语表达的能力,更在于他们能对整个交际事件实施一种反思性的、二阶层面的判断:他们决定目的,而不仅限于优化手段;将“恰当性”视为嵌入具体情境、充满价值负载的判断,而非纯粹的技术指标;能够依据文类、受众与风险识别并纠正机器生成的错误;并将保密、尊严与伤害最小化等约束视为合格实践的内在构成条件。从这一立场出发,核心任务并非仅是人类适应人工智能的效率,而是塑造其实施过程的评价框架、组织框架与监管框架,从而防止“效率”凌驾于准确性、权利与职业完整性之上成为翻译活动的唯一导向。

可以合理认为,当前人工智能系统缺乏意识、意向性和生活经验,而这些要素在需要创造性、默会性文化知识以及可问责性解释的场合中具有重要作用。在翻译活动中,这种缺失之所以关键,是因为即便字面准确或表面流畅的译文,仍可能在细微含义、历史回声、社会定位或文化暗示等层面上有所不足。与此同时,职业翻译领域也完全有理由在某些环节将人工智能视作一种实用工具。译者越来越多地借助机器系统来起草初稿、探索术语、进行表达方案的头脑风暴,以及提升文本的形式一致性(王华树等,2021:9)。

承认这些用途,并不意味着我们要对人工智能不加批判地全盘接受,而是要求我们更精确地界定:哪些任务可以交由机器处理,哪些则必须置于人类的掌控之下。人工智能能够在初步翻译、短语建议、术语管理和重复性内容处理等方面提供辅助;关键在于,最终判

断、语境验证,以及对高风险意义的责任,必须依然由人类承担。由此,有效整合人工智能不仅仅依赖于获得强大的模型,更关键的是建立相应的工作流程:在关键决策环节嵌入人工监督;通过界面呈现不确定性而非将其掩盖;对敏感材料实施严格的数据治理协议;并确保输出结果能够追溯到明确可辨的系统、版本与提示词。真正相关的问题并非人工智能能否加速生产——它显然可以——而是在何种条件下,这种加速依然与职业责任的要求相容。

4 创造力、文学性与高风险文本

人类创造力在翻译活动中仍处于核心地位。毫无疑问,翻译中的创造力并非无拘束的创造,而是一种受规范制约、嵌入具体情境且需承担责任的判断过程。译者在跨越语言传递意义、语气、语力及效果的同时,必须始终关注原文意图、读者期待、文本类型规范以及相关制度或合同要求。当然,译者可以在措辞、隐喻、语域乃至使文本在目标语言中获得生命力的文化替代等方面展现创造性——但这种创造性始终在约束下运作。每一个选择都需在基于忠实与目的的原则之下得以辩护,这意味着译者必须能够阐明为何选择此译法而非彼译法(Nord, 1997)。由此可见,翻译本质上是一种立足于职业责任的“受约束的创造”。译者需要权衡各种竞争性的优先事项,与利益相关方沟通协商,消除歧义,并对最终译文负责。这些活动无法简化为统计模式的简单延伸,而是依赖于具身化的文化知识、长期积累的诠释经验、伦理判断能力,以及对客户、审校者、机构与读者持续负责的职业态度。

当前的人工智能系统尚无法以稳健的方式复现这种受限的创造力。基础模型缺乏驱动翻译决策的社会严肃性所需的动机、价值观念以及对现实世界的实际参与。尽管它们能够模仿与创造性变化相关的模式,却并未真正融入使翻译选择具备辩护性或不可辩护性的实践与伦理环境之中。人类译者能够阐释为何某一译法优于其他方案,说明其中的取舍权衡,依据任务要求的变更进行修订,并在结果不充分时承担相应责任。译者常常需要在动态变化的指令与制度约束下,同时权衡法律风险、安全影响、声誉问题及受众预期(Koskinen, 2008; Pym, 2015)。他们不仅产出文本,更承担着管理交际后果的职责。相比之下,现有模型既无法在法律层面,也无法在伦理层面被追责,其表面上的“解释”亦无法构成任何职业意义上的负责任论证。无论这类系统作为工具多么有效,当前它们仍缺乏被视为真正翻译主体所必须具备的问责能力。

这些局限在翻译无法脱离解释与审美判断的领域中表现得最为明显——也就是文学作品。如果说可问责的决策在一般翻译中已经很重要,那么当意义与风格、歧义以及文本的形式质地紧密纠缠时,这一点就更加重要。自动化的边界在文学翻译中尤为清晰,因为

在这里,人类能动性不是偶然因素,而是构成性条件。文学译者必须处理声音、节奏、歧义、典故、文化共鸣与审美风险,而这些都不是任何模型能够稳定形式化的。任务远远超出短语层面的对等,它涉及语气、节奏、句法结构、互文回响,以及文本究竟应被多大程度地“归化”,或保留多少异质性于目标语之中(Chesterman, 1997)。

对于特定文化负载表达,究竟应采取解释、适应性改写还是原样保留的策略,使读者直面文化他异性而非获得透明化理解,并无普遍适用的先验规则。这一判断取决于文本自身的性质、目标读者群体、编辑背景、出版社的意图,以及译者对文本功能与效果的把握。若翻译仅传递指称意义,却磨平了原文的风格、节奏或文本肌理,其缺陷往往只有经过专业训练的读者才能察觉。西蓬科斯基(Siponkoski, 2013)指出,译者必须在多方利益相关者之间进行协商,其选择也深受这些行动者之间权力关系的影响。实践中,译者通常对审校意见仍保有最终决定权,因为审校工作很少在初稿提交后即告结束。当修订意见出现分歧时,多数译者表示,他们很少接受自己不同意的修改;相反,他们的让步往往具有高度的策略性。遵循“选择战场”的原则,译者可能接受次要的措辞调整——例如某个同义词的替换——以换取在更重要风格要素上的协商空间,例如全文的语域层级或某一角色的独特语气。此类工作倚赖于文学敏感度、文化积淀与说理能力,而这些是当前任何模型都难以真正具备的(Moe et al., 2021)。

与此同时,不能否认在某些领域或环节,机器已经在一定程度上取代了人工。部分替代确实已发生,否认这一点是徒劳的。在诸多常规性、范式化或高吞吐量的翻译任务中,机器翻译与大规模语言模型已显示出比人工翻译更快、更经济的优势。产品说明、界面文本、重复性客服对话等标准化内容,正越来越多地交由自动化系统处理,这也使得人机协作翻译成为业界实践与学术探讨的关键议题。正如Chen(2025: 9)所指出的:“在翻译领域ChatGPT、DeepL等大规模预训练语言模型已在广泛任务中逐渐取代人类译员。”由此可见,该领域所面临的并非一种假设性的未来,而是一个已不均衡展开的现实:自动化正在重新配置翻译劳动的格局。从实践层面看,这一转变正在将人类译员推向当前机器仍难以胜任的工作范畴,例如译后编辑、领域敏感的文本修订、文化语境复杂的本地化、术语体系监督,以及在不确定条件下的质量保障。尽管对简单、常规转换的需求可能会收缩,但随着自动化范围的扩大,对文化中介、学科专业知识、批判性审校与伦理监督的需求反而愈加凸显。

5 未来的关键问题

由此可见,翻译领域面临的核心问题并非是否应全盘接受或彻底拒绝人工智能,而是在一个日益被计算系统结构化的环境中,译者与口译员如何保持真正具有意义的专业能动性。这正是人文学科不可替代的价值所在。正如泰格马克(Tegmark, 2017: 33)所言,当机

器开始涉足那些长期被视为“人类独有”的领域时,问题便超越了技术能力的范畴,转而涉及“人类珍视什么、如何理解自身,以及希望如何组织集体生活”的根本议题。他强调,我们必须反思:人类是否将被这些系统取代、与之融合,抑或仅仅与之共存?而最终,我们应如何在新的时代中主动塑造自身的未来?他进一步指出:“这意味着,如果人类希望明智地引导人工智能的发展,我们不仅需要应对传统的计算问题,还必须直面哲学中一些最为棘手的难题。”(2017: 279)

翻译为这些更广泛的哲学问题提供了一个理想的试验场。翻译从来都不是词汇的机械转移,而是一种深具人文特质的实践,它处于文化认同、伦理判断与制度权力的交汇点。当今翻译行业面临的核心问题,早已超越了技术准确性的范畴——机器能够生成流畅文本已是不争的事实。真正值得深思的是:当跨文化交流中最关键、最复杂的部分被委托给算法时,全球交往格局将发生何种转变?当我们把最敏感的互动交予自动化系统处理时,实际上是在依赖一类工具,而它们从根本上缺乏为其所产生的意义承担道德责任的能力(Kenny, 2022; Moorkens, 2022)。

译文类型是决定任何翻译任务的核心要素。不同翻译需求所提出的要求各不相同,在一场严肃的人工智能讨论中,必须避免将这些任务混为一谈。产品标签与诗歌所面临的翻译挑战并不相同;法庭问答与广告口号在恰当性与风险标准上也遵循不同的原则。哲学、文学、法律、医学以及涉及政治敏感性的文本,均要求超越常规语言转换的审慎判断。在这些情境中,重要的不仅是词汇的准确对应,更在于诠释的分寸感、历史与文化知识的把握,以及识别某些措辞选择所带来的影响已超出形式流畅度所能涵盖的范围的能力(Cronin, 2013; Pym, 2015)。在人工智能介入的条件下,翻译实践越是呈现多样性,就越有必要区分不同的文本类型、制度情境与后果层级,而不能抽象地讨论“替代”问题。

正是基于这一认识,本文的目的既不在于通过含混诉诸“真实性”来为人类译者辩护,也不在于将人工智能颂扬为该领域无可回避的未来。相反,本文聚焦于在技术中介的条件下,翻译与口译中的人类维度——涵盖创造力、伦理责任、情境判断、教育以及专业能动性——如何面临挑战、被重新配置,并在某些情境下甚至得以进一步强化。基于此,本文介入翻译研究、口译研究与人工智能研究中的主要理论争议,并借鉴克罗宁(Cronin)、赫尔曼斯(Hermans)、偌德(Nord)、皮姆(Pym)等学者关于中介、伦理、作者性及翻译行动的理论资源。特别值得持续关注的是人工智能辅助翻译与口译所带来的规范性与教学性后果:我们不仅需探讨实践层面的变迁,更应深入追问,若人类从业者仍需对交际结果保持实质性权威,那么究竟需要何种理论依据、何种培养路径以及何种制度安排予以支撑?

一个重要的问题是:责任必须由个人译者承担吗?实际上,责任也可以由机构承担——比如企业、医院、法院或国家机关,尤其是在涉及机器辅助输出的法律责任上,还应

配备质量评估、审计追踪和程序保障措施。由此可以看出,人工智能并没有消除责任,而是将责任重新分配。此观点值得充分重视。但必须强调,制度框架不能自动运作。在关键决策点上,仍需由人判断某个输出是否符合用途、一种表面合理的翻译是否在具体语境中存在实质性错误、风险是否在可接受范围内,以及是否超越了任何程序保障无法弥补的伦理边界。治理结构可以安排责任的分配,但无法取代责任所需的解释和伦理判断。这些工作依然依赖于人为判断;这不是信念的表述,而是目前所有系统尚未真正解决的实践现实。

6 结语

人工智能对翻译与口译领域所带来的挑战,远不止于技术层面的效率提升或文本生成能力的增强,更是对翻译实践之根本性质的一次深刻追问。机器已能在多种场景下生成流畅、连贯乃至极具说服力的译文,这意味着围绕“机器能否翻译”的初级争论已显不足。真正的议题在于,当翻译不再仅仅是形式的转换,而是深度嵌入法律责任、文化差异、制度权力与伦理后果之中时,究竟谁应对意义负责?谁能在交际结果出现偏差、伤害或误导时,作出可辩护的判断并承担相应责任?显而易见,人工智能正在重塑翻译活动的组织方式,使翻译能动性日益呈现为一种分布于人、机器、平台、流程与制度之间的协作结构。然而,翻译决策中最关键的规范性维度——即解释、判断、权衡与问责——依然无法脱离人类主体而独立存在。人类译者与口译员的重要性,不仅在于他们比机器更具温度或更富创造力,更在于他们能在不确定、冲突与高风险的情境中,对意义进行具体把握,并对由此产生的后果承担责任。正因如此,“人的触感”不应被视作一种模糊的浪漫化修辞,而应被理解为一种专业能力:它涵盖识别语气与语域的细微变化、理解文化与制度语境、处理价值冲突、回应不同利益相关者的诉求,以及在必要时为翻译选择提供理据并承担后果。

与此同时,本文并不主张以简单的二元对立方式拒斥人工智能。恰恰相反,人工智能在术语管理、初稿生成、重复内容处理与形式一致性维护等方面,已展现出显著的实用价值,并在大量标准化、高吞吐量的任务中切实改变了劳动分配格局。问题不在于是否使用人工智能,而在于如何运用,以及在何种制度、教育与职业规范之下运用。未来翻译实践的关键,并非将人类译者塑造为机器无法替代的神秘主体,而是明确界定哪些环节可安全委托于机器,哪些环节必须保留在人类的监督、修正与裁决之下,并通过工作流程设计、风险分级、数据治理与责任制度,确保自动化不会以效率之名侵蚀准确性、权利保障与专业伦理。尤其在文学、法律、医疗、哲学及政治敏感文本等高复杂度或高风险领域,人工智能的局限尤为突出。这些领域所要求的不仅是语义对应,更是对风格、歧义、文化回响、历史脉络、制度后果与伦理边界的综合判断。机器能够模拟表达,却无法真正占据使这些判断具备正当性的社会与伦理位置;它可以生成文本,却不能对文本所引发的影响承担责任。所

以,在这些文本的翻译中,人类译者的作用并非自动化的残余,而是翻译得以成立的根本条件。

由此可见,人工智能时代翻译研究与实践最迫切的任务,并非继续滞留于“人是否会被取代”的抽象追问,而应转向一个更为具体且更具建设性的议题:如何在人机协作日益深化的现实中,重新定义、维护并强化人的专业权威。这意味着翻译教育需培养的不仅是语言能力,还应包括风险意识、伦理判断、技术素养与制度理解;翻译理论需更深入地阐明责任如何在复杂系统中分配而不被消解;而翻译行业与公共机构则需建立能够切实支持人类监督与问责的组织框架。归根结底,翻译之所以始终是一个深具人文意义的课题,正因为它关涉的从来不只是语言,更关乎人与人之间如何理解彼此、如何承担彼此话语的后果,以及如何在差异中建立可信赖的联系。人工智能可以参与这一进程,甚至在某些方面极大提升效率与可及性;但只要翻译仍涉及判断、价值、风险与责任,它便不可能完全脱离人类主体而独立成立。未来真正值得追求的,并非一个无人翻译的世界,而是一个在技术进步中依然保有解释深度、伦理敏感与责任结构的翻译生态。

参考文献:

- 王华树,王鑫. 2021. 人工智能时代的翻译技术研究:应用场景、现存问题与趋势展望[J]. 外国语文(1): 9-17.
- Bender, Emily M. & Alexander Koller. 2020. Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data[C]// *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185-5198.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major & Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? [C]// *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623.
- Chen, Ziwei. 2025. AI Translation and Data Ownership: Ethical Conflicts, Legal Risks, and Translator Agency in the Algorithmic Era[J]. *The Frontiers of Society, Science and Technology*(7): 9-14.
- Chesterman, Andrew. 1997. *Memes of Translation: The Spread of Ideas in Translation Theory* [M]. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Chesterman, Andrew. 2001. Proposal for a Hieronymic Oath[J]. *The Translator*(2): 139-154.
- Cronin, Michael. 2013. *Translation in the Digital Age* [M]. London & New York: Routledge.
- Inghilleri, Moira. 2012. *Interpreting Justice: Ethics, Politics and Language* [M]. London & New York: Routledge.
- Kenny, Dorothy. 2022. Human and Machine Translation[G]// Dorothy Kenny. *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*. Berlin: Language Science Press, 23-49.
- Koskinen, Kaisa. 2008. *Translating Institutions: An Ethnographic Study of EU Translation* [M]. London & New York: Routledge.
- Moe, Marija Zlatnar, Tamara Mikolic Juznic & Tanja Zigon. 2021. Who Determines the Final Version? The Roles of Translators, Language revisers and Editors in the Publishing of a Literary Translation? [J]. *Across Languages and Cultures*(1): 14-44.
- Moorkens, Joss. 2022. Ethics and Machine Translation[G]// Dorothy Kenny. *Machine Translation for Everyone: Empowering*

- Users in the Age of Artificial Intelligence*. Berlin: Language Science Press, 121-140.
- Nord, Christiane. 1997. *Translating as a Purposeful Activity: Functionalist Approaches Explained* [M]. Manchester: St. Jerome Publishing.
- O'Brien, Sharon. 2012. Translation as Human-Computer Interaction [J]. *Translation Spaces*(1): 101-122.
- Olohan, Maeve. 2011. Translators and Translation Technology: The Dance of Agency[J]. *Translation Studies*(3): 342-357.
- Pym, Anthony. 2012. *On Translator Ethics: Principles for Mediation Between Cultures* [M]. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Pym, Anthony. 2015. Translating as Risk Management[J]. *Journal of Pragmatics*(85): 67-80.
- Siponkoski, Nestori. 2013. Translators as Negotiators: A Case Study on the Editing Process Related to Contemporary Finnish Translation of Shakespeare[J]. *New Voices in Translation Studies*(9): 20-37.
- Tegmark, Max. 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence*[M]. New York: Alfred A. Knopf.

Human Agency and Accountability in AI-Assisted Translation

SUN Yifeng

Abstract: This article examines how AI is reshaping translation as a practice, a profession, and a site of responsibility. The central question is no longer whether machines can produce fluent target-language text but who is accountable when translations carry legal, ethical, cultural, or institutional consequences. Translation, on this account, is not linguistic substitution; it is a mediating practice embedded in institutional contexts and dependent on professional judgment. The article views AI's role in translation as genuinely double-edged. Machine translation and large language models offer real gains in terminology management, draft production, and the handling of repetitive or standardized texts — gains that are already restructuring how translation labor is organized. But the fluency of machine-generated output can mask semantic bias, tonal misalignment, cultural incongruity, and institutional risk. What this paper calls the “human trace” — the translator's capacity to interpret, evaluate, revise, and take responsibility in context — becomes more necessary, not less, as machine output grows more convincing. AI does not remove the need for human translators. It redistributes translational agency across a human-machine workflow in which normative judgment and final accountability must remain with the translator.

Key words: AI translation; translator agency; translational accountability; human-machine collaboration; ethical judgment

责任编辑:龙丹