

第六章

伦理道德与机器翻译

Joss Moorkens
都柏林城市大学

神经机器翻译比以往的机器翻译范式更利于促进交流，但也会导致一些后果。和任何技术的发展一样，机器翻译并非道德中立，而是反映了其背后开发者的价值观。本章围绕机器翻译的伦理问题，以数据的收集和重复使用为切入点，探究机器翻译如何才能契合译者的价值观和准则。若能反映价值体系，机器和系统能明确地“好”，并消除译文中的偏见吗？机器翻译对可持续性和多样性的讨论有何贡献？本章的目的不是推广不假思索地遵循指令来运用技术的方法，而是强调在设计数据驱动的机器翻译工作流程时，有意识的决策过程的重要性。

1 何为伦理规范？

伦理学领域研究道德规范、善恶对错，并回答如何才能更好地生活。对这类主题的讨论最早可以追溯到埃及、巴比伦和印度。苏格拉底等希腊哲学家引入了“美好生活”（good life）的概念，这是一种有价值且令人向往的生活。亚里士多德提出了一套具化这种生活的道德准则，人类如果遵照这些准则，就能繁荣发展。不过，这些准则还是比较抽象，不总能帮助我们判断行为的对错。之后的哲学家和伦理学家也提出了一些判断行为是否正确或合乎道德的方法，例如，是否能为大多数人谋福利或只出于良好纯粹的动机。

而问题在于，对某个群体来说动机善良或带来益处的事情，并不一定能给另一个群体带来同样的正面效果。对此，两派观点争执不下。有人认为，行动可以是普遍有益且合乎道德的，如维护正义或诚实，但还有人认为，价值观可能因个人或群体而异，这两种观点之间存在矛盾。关于如何根据主体性、关系或周边叙述来理清某个行为是否合乎道德，已经有很多这方面的建议。理论伦理学就此迈入应用伦理学，试图为我们提供具体场景的行动指导。

在职场领域，应用伦理学通常包括指导职业行为的准则或标准。这些准则如果过于严格，可能阻碍潜在发展或社会效益。僵化的准则还可能导致陷入



困境，因为道德决策很少是非黑即白的，选择可能受到决策者所面临的特殊情况和压力的影响。正因如此，不同领域的应用伦理学便应运而生，根据特定职业场景来思考共同的问题和困境。本章将围绕与机器翻译关系最密切的领域，如计算机和信息伦理、数据伦理，讨论人类在系统开发过程中，如何合乎道德地使用机器翻译。本章第三节节 3 借鉴商业伦理和越来越多的翻译伦理相关研究，阐述了如何在翻译工作流程中合乎道德地使用机器翻译。第四节节 4 借鉴机器伦理和计算机与信息伦理，探讨计算机如何成为伦理主体。而最后几个小节则借鉴各项最新相关研究，聚焦伦理和人工智能的可持续性和多样性。

技术逐渐成为人类生活必不可少的一部分，且计算机的运用也进入到很多领域，因此技术领域的道德准则也受到越来越多的关注。在设计、开发或使用技术的过程中，我们需要意识到自己所做选择的影响。常有人认为，技术是道德中立的，只在人类使用该技术时才会产生偏见。然而，伦理学家和科学哲学家都一致认为，技术并非道德中立，而是反映了设计者的价值观。这些价值会决定技术要解决的问题、创造技术的决策、落实方法、预期用户、所使用的参考或训练数据、数据处理、数据存储的地点和安全性，以及可能基于成本或地理位置的技术获取限制。

科技发展之快且规模之大，都让监管措施无可避免地稍有滞后，因此，我们只好依靠工程师和开发商能遵照道德准则行事。我们也或多或少依赖有政治影响力和财力的大型科技公司能以人类集体的最大利益来行事。但近年来的一系列报告和披露表明，这些科技巨头辜负了公众的信任。技术虽然有助于开辟途径和增加收益，但也会导致公众面临风险。本章将讨论机器翻译系统开发和使用过程中的选择和风险，旨在指引用户做出明智且合乎伦理道德的决定。贯穿始终的重点主要（但不完全）是以传播为目的的机器翻译，即机器译文并非生产流程的最后一步。

2 机器翻译系统开发：人类合乎伦理道德的机器翻译使用

2.1 数据使用的案例研究

本节着重探讨机器翻译系统开发过程中，翻译数据使用的法律和道德问题。正如Pérez-Ortiz et al. (2025 [本卷])所证，数据驱动的机器翻译，尤其是神经机器翻译，需要大量数据进行训练。¹ 尽管重复使用翻译数据来训练机器翻译系统完全合法，但这是否合乎道德呢？思考以下例子，或许有利于理解本节要讨论的问题。

¹ “数据”指任何形式的记录信息，通常以数字形式存储。我们将可大量获取并规模处理的数据，称为“大数据”。

译者 A 自愿与经常合作的客户签署了一份合同,以自由职业者的身份,在专门的网络平台完成翻译任务,并明确准许其翻译数据可用于训练机器翻译系统。该雇主使用译者 A 等人的译文来训练机器翻译系统。随着时间的推移,译者 A 的语言对的神经机器翻译系统质量得到了提高,于是,该公司将翻译任务改为译后编辑,并单方面把按词计算的翻译费用降低了 30%。之所以按此幅度下调,是因为该公司通过平台收集到的翻译活动数据获知,翻译效率提高了约 30%。为了增加收入,该公司决定对外出售机器翻译服务,其中还包括为武器制造商提供的翻译内容。

译者 B 出于个人原则而反对机器翻译。该译者的合作公司希望译者提交包含目标文本的翻译记忆库,以便日后供其他译者使用。但是,译者 B 并不知道自己的译文已经由自动化项目管理系统自动分配,既没有翻译任务说明,也没有与公司直接沟通。此外,译者 B 也不知道,该公司不久后会被一家大型企业收购,该企业会用所有手头数据来训练机器翻译系统,然后出售。数据在共享前应删除所有个人信息(参见节 2.5),但当数据传给某个买家时,这些信息都意外保留了。该公司试图不做声张,以逃避责任。

就以上两个案例,你能从中发现哪些道德问题?如果雇主或译者做出不同的道德决定,会出现什么变化?我们会在下文探讨数据的所有权、使用权限、流通、隐私,以及数据共享的法律框架。希望这些内容能引导你对上述问题的思考。

2.2 数据的所有权

数据常被比作石油,让人误以为大数据是自然形成的,而实际上,它是由人类创造而成。近年来,数据呈指数级增长,这意味着如今无论是可用于训练机器翻译系统的数据,还是对翻译的需求,都比以往任何时候都多,远远超过了人工翻译的生产能力。机器翻译训练数据通常以人工翻译的平行或对齐的双语文本语段的形式,存储在翻译记忆库中(尽管机器翻译译文有时也用于系统训练)。这些翻译数据可能来源于公共数据库(如欧盟委员会总司翻译数据,该数据库可“出于商业或非商业目的,免费复制和传播……”,可参见Steinberger et al. 2012: 457)、个人的翻译数据库或从网上爬取的平行数据。

1886 年首次颁布的《伯尔尼公约》为翻译版权提供了法律基础。根据《公约》,译文属于衍生作品,“应作为原创作品受到保护,但不损害原创作品的版权”(World Intellectual Property Organization 1979 第二条)。《公约》赋予原创作品的作者授权翻译的专有权,但并没有清晰界定何为“原创作品”或“原创性”,导致不同管辖权对此有不同的解释。Troussel & Debussche (2014)认为,“创造性翻译”可视为具有原创性,不过这还有待法庭检验。他们还认为,根据欧洲数据库指令,如果“对内容的获取、查证或呈现进行了大量投资”,就可以主张对翻译记忆库的所有权(European Parliament 1996 第七条)。但在

实践中,无论译者是否在合同中约定放弃记忆库所有权,翻译记忆库通常都会发送给客户。

在大规模范围内,翻译大数据已经成为机器翻译和机器学习系统训练的宝贵资源 (Moorkens & Lewis 2019a)。然而,这并不意味着要降低译者的报酬,而且机器翻译训练中的数据细粒化重复使用意味着,训练数据通常无法溯源。通过网页爬取收集平行文本数据也是如此。在我们的案例研究中,译者 A 和 B 除了交出自己的翻译数据并接受随之而来的后果之外,别无选择,特别是大多数译者都是自由职业者,没有太多雇主争论的余地。我们有理由认为,更公平的数据所有权制度会有助于翻译行业的可持续发展(还可参见节 5)。

2.3 数据的使用权限

在一些司法管辖区域,支付费用的雇主被认为是翻译成果的合法所有人,而在另一些区域,所有权可以转让,并准许他人重复使用。我们认为,那些愿意与其他译者,甚至陌生人,共享资源的译者之间存在“一定程度的协作关系”(Moorkens & Lewis 2019b: 8)。然而,若翻译数据用于训练机器翻译系统,译者对自己的翻译数据重复利用的接受度就会降低,尤其对于那些认为机器翻译技术的进步会损害人工翻译利益的译者来说。一些翻译合同可能明确规定,翻译数据会用于人工或机器翻译工作流,但少有译者能左右自己的译文会如何被重复使用。同样地,通过爬虫获取的互联网数据也从未考虑到使用权的问题。

这意味着译者 A 和 B 对今后的项目有所贡献,但他们并不了解这些项目的目的和最终用途。译者 A 可能出于道德因素,反对为武器制造商工作,但仍会在不知情的情况下参与其中。对于数据被搜集并重复使用,以及为大型技术项目做贡献的人来说,这种不透明性更加普遍存在,他们并没有机会询问:“我的工作的最终应用和用途是什么?”以及“我会因参与实现该用途感到满意,还是羞愧?”(参见 Moorkens 2020 和 Weizenbaum 1986)。

翻译过程的数据创建,会根据保存和交换格式记录下许多属性,通常包括译者的姓名或 ID (参见节 2.5)、创建日期和时间、语言代码、使用的软件和项目 ID。这些信息不仅有助于按照译者、项目或创建日期来决定何时何地重用数据,还可以记录译者的翻译任务数据,包括更详细的计时、编辑动作和每一次键盘活动记录,特别是译者 A 在项目中使用的专门网络翻译平台。这些数据可用于监督译者的工作,但通常在训练机器翻译系统时删除,只保留平行句对。一旦元数据(关于数据的数据)被删除,就无法保存未来使用或重复使用的偏好,也就无法衡量个人贡献,即使协议有追溯性变更,这意味着贡献应获得版税。另一方面,抹去元数据将改善翻译数据的匿名化,该操作对于数据的共享或交换来说非常重要。

2.4 数据的流通

Topping (2000)写道,在翻译记忆共享的早期阶段,独立译者、本地化机构和本地化客户之间共享翻译记忆库,只有译者认为这种做法合乎道德规范。而到了 2021 年,这种观点似乎变得怪诞,因为有些公司的商业模式就是基于搜集并转售为机器翻译和其他机器学习技术所用的数据。²这种商业模式切实可行,因为训练数据量和数据管理水平都会影响到 NMT 系统的质量和价值。

正如节 2.2 所提到的,某些数据集可以通过开放平行语料库 Opus 等项目免费提供和流通 (<http://opus.nlpl.eu>, Tiedemann 2012),因为这些数据集都有许可协议,或受到欧盟关于公共部门信息再利用指令 (2003/98/EC) 的保护。不过,这并不一定意味着译者明确授权所有可能的重复利用形式,但他们知道数据将与公众共享。

数据可能会根据公司之间的协议,或因某公司被其他公司收购而流通,这在语言服务行业中司空见惯 (Moorkens 2020)。这些数据可出于科研或慈善目的被买卖或捐赠,而无需事先获得数据创建者的批准。这么做完全合法,除非涉及到限制使用的个人数据。

2.5 私隐与个人数据

可识别个人身份的数据被视为个人数据,其中包括含有创建者姓名或编码(匿名)的翻译记忆库。在欧盟内部,《一般数据保护条例》(General Data Protection Regulation, GDPR)³自 2018 年起限制个人数据的共享和再利用,为国家立法提供了指导方针,对数据泄露行为予以高额罚款。这不仅提高了网络安全,还限制了服务器在非欧盟地区的使用。除研究目的等例外情况,任何个人数据只有初次使用许可包括允许再利用,才可以再次利用。欧盟以外,还有许多其他国家出台法规管理个人数据的使用。

公司应该报告任何数据泄露情况,但有不少揭露数据泄露事件的媒体报道反而被公司压了下来。原因之一是为了逃避 GDPR 规定的罚款(根据该条例第 83 条,罚款最高可达 2000 万欧元或年营业额的 4%,以金额较大者为准),但除此以外,公司也希望能避免影响企业形象、失去消费者的信赖,且如果是上市公司,还需要担心股价下跌的问题。理想情况下,透明度会提高公众对企业的信任度,但大型企业的数据安全问题并非总是清晰明了,其庞大规模使得数据保护颇有难度。此外,并非所有数据泄露的情况都一样,有可能是受聘于公司的“道德黑客”(ethical hackers)为了识别漏洞以保护网络安全或公众,也有可能是恶意黑客为了谋取私利而窃取数据。

²有关机器学习的更多信息,请参见本书的导言。就本章而言,我们对机器学习的理解是,使用计算机借助大数据进行推理,而非根据输入的明确命令来达到目的。

³欧盟法规 2016/679, 获取连接<https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN>

一旦清除个人属性并匿名化,那这些就仅仅是可分享的数据而已。⁴当然,即使没有元数据,如果内容或风格容易辨认或生物特征数据与个人有所关联,那么这些也是可识别的数据。如果数据被共享或汇集,来自个人或群体的数据可用于推断某些信息,如种族或性别等受 GDPR 特别保护的属性。由于这些推论基于综合数据得出,而非明确包含在任何单一数据集中,因此它们通常不在 GDPR 的范围内 (Wachter & Mittelstadt 2019)。这给“群体隐私”(group privacy)造成了风险,某个群体会因无法识别个人的数据内容而受到歧视 (Floridi & Taddeo 2016)。

例如,译者 A 和 B 可以汇集自己和他人的翻译数据,准许第三方对他们个人或作为群体的一员做出相关推断。随着在线翻译平台的使用越来越普遍,译者不仅更无法掌控其翻译数据,还要允许翻译过程受到监控。翻译平台会收集翻译活动数据,若平台捕捉到个人情况导致翻译效率或质量出现短暂下降,这可能会对他们未来的就业前景产生负面影响。如果这些记录翻译表现不佳的、可识别个人信息的数据被泄露出去,可能会对该译者产生深远的不良影响。这并不一定意味着监控质量或生产率的做法不符合道德规范。翻译中介或公司需要能够监督他们的翻译。然而,公司将雇佣决策或沟通过程自动化后,如译者 B 的项目管理,译者便没有机会解释翻译决策或建立长期的信任关系。即使在最理想的情况下,也无法保证译者或平台用户的行为符合道德规范,但研究表明,若只是纯粹的交易关系,双方的信任和假定诚信尤其会受到破坏,从而产生连带效应,影响满意度和绩效 (Whipple et al. 2010)。

2.6 机器翻译评测中的伦理问题

Rossi & Carré (2025 [本卷]) 探究了机器翻译质量的人工和自动评测方法,有许多与此相关的道德规范问题值得在此提及。大多数机器翻译系统会在训练过程中对译文进行自动评测,待训练完成后,以简单快速的低成本方式再次进行评测。在机器翻译共享任务竞赛中,开发团队带着各自的机翻系统一较高下,系统性能的衡量方法包括自动评测或大众评分。基于这些评测,如果句段级别的群体评分或自动评测得分与人工“参考译文”的相一致,那么机器译文就被认为已经能与人工翻译相媲美;如果机器译文的得分超过人工翻译,那么前者就被认为是“超越人工翻译”的译文。

这种说法值得商榷,尤其涉及到传播范围更广泛的学术出版物和市场营销文案。甚至新闻媒体还会报道这种观点,让人觉得机器译文完美无缺,毫无风险,不再需要人工翻译。然而,自动和人工评测结果相关性极其微弱,且众包评分存在几个问题,如对连贯的译文语段进行排名或打分的是匿名或未经培训的互联网用户。Freitag et al. (2021)发现,能获得完整原文和译文的

⁴在撰写本文时,有很多人正致力于实现翻译数据的自动匿名化,但很难取得可靠的结果。

专家(职业译者)评测人员在详细的翻译错误分析时,得到的结果与众包评分者的存在显著差异,且明显更青睐人工译文,而非机器译文。此外,报酬低廉、工作条件差、用户评级系统不透明,以及在无监督或道德审查的情况下利用人类(众包工作者)参与研究,这些都是众包形式的问题所在。尽管如此,基于自动评测和众包评分的测评结果大多也只是简略提及,既贬低了人工翻译的价值,还让翻译客户等大众对机器翻译产生了不切实际和不加批判的看法。这种看法使得机器翻译更有可能被纳入翻译工作流程。

3 翻译工作流程中合乎道德规范的机器翻译使用

3.1 翻译的利益相关者

节 2.1中的案例研究不仅引发了我们对伦理和机器翻译系统开发的思考,同时也促使我们思考如何在翻译工作流程中,合乎道德规范地使用机器翻译。例如,雇佣译者 A 的公司单方面决定将翻译工作改为译后编辑,但理想情况下,将机器翻译嵌入翻译工作流程的前提是通过协商,得到所有利益相关者的一致同意。如节 2.1的案例所示,利益相关者包括翻译代理机构或语言服务提供商,其中还包括多个内部角色,以及自由职业译者。此外,翻译客户应知晓翻译过程中会使用机器翻译,并应了解相关的利弊(参见节 3.2)。几乎所有关于译后编辑翻译效率的研究都显示,与从头开始翻译或使用翻译记忆库相比,译后编辑的翻译效率更高。然而,一般认为,机器翻译的使用应该与翻译文本的可用时限和风险程度相关。翻译客户依赖于中介的专业能力来选择合适且性价比高的翻译流程,而终端用户则依赖于翻译客户提供可靠的目标文本。此外,Pym (2012)提出,即使低风险目标文本的机器翻译质量差,难以理解,但这种译文也有可能符合客户需求(通过轻度或完全译后编辑来降低成本),不过,最终读者需要花费更多精力理解文本。

在使用机器翻译的工作流程中,翻译软件开发人员也是重要的利益相关者,节 1中提到的技术开发人员的价值和相关设计决策也适用于这一群体。所使用的翻译工具可以改变翻译活动及其意义。与机器翻译的互动可通过交互式机器翻译进行,其中机器翻译用于自动建议、动态编辑或译后编辑,作为额外的翻译建议或自动扩展显示在译文窗口。翻译工具注重使用性,例如兼容惯用的键盘快捷键、提供整洁的界面,以及尽可能多的个性化设置。翻译活动数据(参见节 3.2)若被收集,可以对用户公开,这样他们就能看到收集内容并自愿使用。译者也可以使用削弱译者自主能力的翻译平台,只会即时接收任务而看不到数据收集情况,交互界面的功能比较有限,且客户可以对译者表现进行评级,但没有提交反馈或讨论的选项。

译者可以根据文本领域和工作条件,来选择接受或拒绝某项翻译工作。但对于那些不了解机器翻译可以通过多种方式融入翻译工作流程的人来说,可能很难做出选择,尤其在翻译中介的信息不透明的情况下。同样重要的是,

译者要说明工具使用情况,尤其是机器翻译,以确保遵守保密协议,不会在无意中损害其他利益相关者的利益(或如节 3.2所述,有损自身利益)。对于译者来说,遵循职业道德准则,即指导合乎道德的行为规则,属于义务伦理学的范畴。尽管这些准则是对翻译中道德角色的狭义理解,但它们仍然在决策中起到了积极的作用。⁵和许多其他行业组织一样,翻译协会通常发布这类行为准则,以鼓励专业操守、公正诚实和遵守保密原则。这些准则还有助于建立与现有和潜在客户之间的信任关系。撰写本文时,我在查阅这些准则的过程中发现,准则并没有直接提及机器翻译,即使(正如我们刚刚讨论的)决定在翻译项目中是否使用机器翻译,也可能涉及到伦理问题。Chesterman (2001)撰写了很多有关翻译道德和信任的论述,提出了一种以服务伦理为重点的通用原则,强调要忠于客户提出的条款和质量要求、源文本及其作者以及目标文本读者。

使用机器翻译不一定会导致翻译质量变差,而且耗时费力的高成本人工翻译也不适于所有类型的文本,特别是那些时限短且风险低的文本。然而,对于误译后果较严重的重要文本来讲,使用机器翻译必须要谨慎考虑和反复审查。有证据表明,译者使用机器翻译时,一些项目经理会睁一只眼闭一只眼(Sakamoto 2019),但翻译客户有理由要了解机器翻译的使用情况,并在合同中明确说明是否可以使用。

3.2 风险和责任

如节 2.2所述,翻译合同可以规定翻译数据的所有权,或允许译者或客户保留翻译记忆库。而Canfora & Ottmann (2020)将在下文介绍有关机器翻译使用的其他两个合同规定范畴——法律责任和保密条款。

译者可能会因为疏忽或未能履行对客户的注意义务而违反合同。责任只能指人的行为,这意味着行为人必须对在翻译过程中,由于机器翻译导致的错误而造成的伤害或损失承担责任。暂且不论责任,机器翻译可能给最终用户带来风险的这一事实就是道德伦理问题。关乎安全的内容不应该使用原始的机器译文。目前的翻译和译后编辑标准没有提到最终用户的责任或风险。

不过,ISO 翻译标准确实强调了“安全和保密地处理所有相关数据和文件”的重要性(ISO 2015第 3.2.a 节)。免费在线机器翻译系统的用户授予服务供应商使用用户输入数据的权利,此外,还存在不经思考而使用免费在线机翻系统来翻译机密和敏感材料,从而导致数据泄露的情况。正因为机翻系统会造成这类网络安全风险,Canfora & Ottmann (2020)才认为,最好选购不会保留数据的机器翻译服务,并建议公司使用服务器架构不对公众开放的封闭平台,或为了保密,根本不把翻译任务外包出去。但正如第 2 节和第 3.1 节所

⁵Lambert (2018)和其他学者认为,许多道德准则强调的翻译中立性假设,与对技术的假设一样存在缺陷。

述,自由译者对封闭平台较为反感,因为无法自行管理翻译数据和译者活动数据。

ISO 译后编辑标准并未提及(ISO 2017)保密性或数据安全风险,这相当令人惊讶,因为译后编辑必然会用到机器翻译,从而可能引入风险。信任是降低风险的关键,因为只有当译者认为自己与合作伙伴相互信任时,标准、指南和合同才有价值。若无信任,他们可能试图合理化不道德的行为(参见Abdallah 2010)。

除了有风险顾虑,由于节 2 中描述的流程或人工智能(AI)对工作领域和可持续性的影响,译者和用户可能并不想使用机器翻译。下一节内容会从神经机器翻译的角度,来深入探讨人工智能对翻译领域和可持续发展的影响。

4 可持续性

4.1 译者/译后编辑者的报酬、条件和工作满意度

虽然翻译任务需要精湛的专业技能,但在译者 A 和译者 B 的例子中,部分翻译流程已经(在一定程度上)实现自动化,如自动分配任务、进行译后编辑,以及翻译数据使用目的发生转变,用于译者不知情的任务。越来越多的人认为,人工智能将对许多以往看来不受自动化影响的工作领域产生重大影响。虽然可能不会直接造成失业率上升,但这些变化可能会以难以预测的方式影响经济收入、工作组织和技能管理。这些在很多行业都是未来的考虑因素,但对翻译领域已经产生影响,原因如下。首先,自 2010 年左右以来,机器翻译的译后编辑一直是翻译市场增长最快的领域,早于从统计机器翻译向神经机器翻译的转变。自翻译记忆库在 20 世纪 90 年代早期出现以来,储存翻译数据的做法已经司空见惯,尽管以监控和自动化为目的的翻译活动数据收集才出现不久。其次,很大一部分译者为自由职业工作者意味着他们兼具灵活性和自主性,但要在一个接一个项目的时间框架内工作。这造成了权力的不平等——在一个个项目中,翻译中介和雇主可以单方面改变的翻译过程和工作条件,但译者对此几乎没有发言权。节 2 关于数据的讨论清楚展现了权力差异的影响。随着合并和收购的步伐加快以及大型翻译上市公司的成立,商业运营决策者与从事翻译、译后编辑、审校、标注、修订、字幕制作或其他与文本直接相关工作的众多自由职业者,两者脱节越来越严重。翻译行业提出将项目管理自动化,以及使用区块链来确定确定著作权,都不太可能改善这种情况。更概括来说,正如 Moorkens & Rocchi (2021)提到的,翻译行业对伦理道德问题一直未予以重视。

早在 1980 年, Kay 曾就翻译技术提出一个观点,即译者应掌握技术,以辅助完成机械重复且单调乏味的工作(Kay 1980)。但我们所见却正相反,许多译者的工作因为技术而受到限制。一些译者的工作内容已经简化为质量检查、标注或改正质量参差不齐的机器翻译中重复出现的错误。在改正机翻错

误时，机翻译文甚至可能切分成无语境的独立句子，有些自动通过质量检查，有些则标记为待审校。

有些译者喜欢进行译后编辑，即便薪酬降低（如节 2 第一个案例研究中的情况），他们也觉得这种工作收益好，值得做。但是，在技术化的翻译过程中，所有利益相关者都需要在短期效率和长期收益之间取得平衡。

从业者的满足感来自有意义的工作、对任务的掌控以及与同事共同奋斗。他们的工作动力是成就感和对这种成就的认可。如果在翻译过程中没有成就感，那更高水平的译者会离开这个行业，从而导致高水平的译者和译后编辑者的短缺。这种短缺会影响多语信息的可靠获取，以及机器翻译训练所依赖的高质量双语数据的收集。Docherty et al. (2008: 4) 认为，可持续的工作系统必须满足许多，而不是少数利益相关者的需求；不能只关注“短期的静态效率，如生产率和盈利能力，还必须注重长期的动态效率，如学习和创新”。联合国可持续发展目标 (UN Sustainable Development Goal) 第 8 点是提供体面的工作和经济增长，但该目标的第 13 点和第 15 点所涉及的环境可持续性也与机器翻译有关。

4.2 环境问题

我们有理由认为，经济增长目标与环境可持续性目标之间难以兼顾。Cronin (2017) 特别指出了本地化行业的增长依赖性 (growth dependency)。机器翻译所依赖的信息与通讯技术 (ICT) 行业需要开采稀有金属，且因资源回收不善和对环境造成污染而久负骂名。神经机器翻译的资源密集特点尤为突出，如需要性能强大的图形处理器 (GPU) 进行训练，且能源消耗极高。Strubell et al. (2019) 估计，训练一个 Transformer 神经网络大模型所产生的二氧化碳排放量，大约是一辆汽车整个生命周期（包括燃料）的 5 倍。⁶ 然而，大多数模型训练的资源密集型程度远低于本书所举的例子。此外，硬件的性能更加强大，设计和生产成本也越来越高，但随着能耗优化，且大量程序能够同时运行，训练所需能耗也相应减少。尽管如此，神经机器翻译系统的训练成本依然高昂，耗电量大。这对环境的影响将取决于能量的来源。发布机器翻译系统的详细信息时，目前还没有商定的能耗基准，尽管在可持续人工智能发展建议的背景下，已经提出了一些基准。Van Wynsberghe (2021) 强调，如果不关注人工智能（由此推及神经机器翻译）开发及运用的可持续性，人工智能开发本身将是不可持续的。

5 多样性

5.1 开发者和用户

开发神经机器翻译的巨大障碍在于成本和能耗。对数据的要求意味着早期的系统只能使用公开可用的数据（参见节 2），通常为主要的欧洲语言创建系

⁶我们还注意到，只有规模最大的公司才负担得起训练这类大模型的费用。

统。因此，无怪乎早期发表的神经机器翻译相关工作主要由北美和欧洲的资源丰富的学术研究小组主导。但这种情况有所改变，原因有二。首先，大型技术公司全力支持神经机器翻译的研究工作，建立了资源丰富的团队，这些团队在优化主要语言对的机器翻译系统方面处于领先地位。这意味着许多学术研究团体难以跻身主要的欧洲语言领域，只能转向低资源语言和少数民族语言系统开发这一“小众”领域。其次，通过将单语数据从预期目标语言通过机器翻译成预期源语言⁷来合成平行数据的能力，提高了低资源语言对的机器翻译质量。因此，第五届国际机器翻译大赛（WMT20）包括因纽特语和泰米尔语与英语的互译任务。然而，提高低资源语言翻译质量的另一种方法是构建大型多语言系统，不过通常是大型团队才有资源和财力这么做。

目前还没有关于机器翻译研究团队多样性的调查。不过，只需要搜索相关学术论文就会发现，有相当数量的论文使用简体中文、孟加拉语和印地语来发表，这些语言的使用者众多，但都不是欧洲语言，且孟加拉语和印地语在历史上一直是低资源语言。在研究团队带头人与研究议程决策者当中，多样性可能较低。当论及机器翻译之外的偏见，这种多样性的缺乏是一个突出的问题，导致了使用机器学习的系统中许多广为人知的错误和盲点，例如为识别疑犯而开发的人脸识别系统过度倾向于将少数族裔划分为罪犯。

Vieira et al. (2020: 13) 在关于机器翻译的社会影响的文章中指出，在用于医疗和法律领域的案例中，机器翻译的使用不当会对弱势群体造成伤害。他们发现，这可能会加剧不平等现象，因为“世界上相对较少的语言拥有的数据和资源有些不成比例”，加之机器翻译的使用者缺乏机器翻译素养。另一方面，机器翻译有助于促进交流，因为越来越多的免费机器翻译系统增加了低资源语言选项。在撰写本文时，谷歌翻译（Google Translate）共涵盖 109 种语言，包括齐切瓦语、苏格兰盖尔语、维吾尔语和鞑靼语等低资源语言。

5.2 机器翻译的输出

Jones 于 2016 年算出，在超过 6000 种非濒危语言中，只有 1% 的语言被机器翻译系统覆盖。得益于节 5.1 提及的研究，这种情况有所改善，百度和谷歌覆盖的语言数量越来越多也证明了这一点。但是，在现实世界中，这发生在一个信息倾向于从资源丰富的语言，流向资源贫乏的语言的世界里。由于机器翻译比人工翻译对源语言的干扰更大（参见 Toral 2019），人们担心从长远来看，资源匮乏的语言会愈加贫瘠。

所有能用机器翻译的语言都可能面临类似的处境，特别是如果人工翻译的数据不足意味着机器翻译系统很难跟上当代语言发展。Vanmassenhove et al. (2019) 展示了，当神经机器翻译引擎的训练至所谓的收敛点⁸，会导致词

⁷机器翻译研究人员称这一过程为“回译”。不要将“回译”与 Baker (2018) 等传统翻译研究资料中，用作注释技术的“反向翻译”混淆。

⁸即自动测评分数显示输出质量没有改善时，迭代的神经机器翻译系统训练停止的点。参见节 7.2。

汇多样性下降,这表明若训练在更早的时候停止,神经机器翻译输出则不那么同质化,且词汇更多样化。Vanmassenhove (2019)表明,这种输出译文的标准化解显出算法偏差,这种偏差加剧了训练数据的性别偏差,其中名词等形式与最常见的某个性别相关联(在两性制度中),而不太常见的性别则在输出数据中被标准化了。这导致输出译文加深社会偏见,性别使用不一致,甚至在单个句段中也是如此。下节内容将讨论目前如何消除输出中的偏见,以及作为伦理问题的隐藏主体,计算机还发挥着哪些作用。

6 伦理主体——计算机

有关人工智能道德问题的讨论主要涉及安全,并将提升机器的自主性以及将人工智能技术的应用拓展到更多任务。与此同时,机器道德问题的特点在于,它将“机器”视为具有“能动性”(即行动的意愿和能力)的主体,而不是客体。关于机器翻译的偏见和风险的研究虽不多,但呈现日益增长的趋势(参见节 5.2),此类研究的范围也囊括更宽泛的机器翻译和人工智能的内在价值观(参见节 5.1)。这便催生了计算机作为伦理主体的概念,因此道德决策隐含于计算机的设计之中。虽然已有这方面的努力正在进行,但我们经常在网上看到有偏见或歧视性的机器译文,其出现场景对于开发者而言难以预测。当机器似乎显露出智能,技术(尤其是人工智能)带来的意外后果启发了一系列探讨机器翻译技术使用有违伦理道德规范的书籍和文章。

机器翻译技术的问题可能涉及数据收集、用于从数据中归纳提取模式的算法,以及数据本身。技术带来的结果也可能是有意为之,也就是说,由此技术的“预设用途”(affordance,即工具所暗示的使用方式和互动方式)推动用户以某种方式行动,而这种方式可能对他们自己或他人带来负面影响。

在关于神经机器翻译偏见问题的少量研究中,研究重点仅限于性别偏见,而非种族或性向(sexuality)等其他属性,并提出了一些解决方案来“消除”机器译文的偏见。在神经机器翻译译文中,如何正确使用性别和消除性别偏见,建议方法包括使用性别标签(类似礼貌用语和语域的标签)、消除偏见的词嵌入,或者用类似领域调试(domain adaptation)的方法处理性别偏见,比如利用小型性别平衡数据集进行迁移学习(Tomalín et al. 2021)。2020年,谷歌针对某些语言推出了一个系统,对每个带性别标记的名词同时输出阴性和阳性两种形式的译文。用户可以自行评估不同的选项,选择合适的译文。虽然这种做法仅基于二元性别论,但还是提高了谷歌翻译处理性别信息的译文质量(Johnson 2020)。

未来,计算机有可能变成显性的伦理主体,有能力处理每种场景的信息,并自主决定最佳或最符合道德规范的行动方案。但目前的技术还未发展到这种先进程度,因此机器创造的内容还不能声明版权,机器也不能对损失或伤害负责。即便技术很成熟了,我们也不能假定计算机将按照道德规范来行

事。我们已经确定技术并非无害。虽然技术的使用能够造福个人和社会,但也存在一些使用之前难以预料的风险。

毫无疑问,免费在线机器翻译工具的使用促进了很多人的交流,但是无缝界面和质量更高的输出可能会让终端用户认为,该技术的使用不会造成伤害(通常不会)。节 2.1 所举的例子中,翻译工作流程的利益相关方可以被认为是具有语言技术方面的专业知识,但对于普通大众来说并非如此,他们可能期望机器翻译具有与(优质)人工翻译相同水平的连贯性和可理解性。他们甚至可能意识不到自己正在读的是机器译文,即使译文被标记为机器翻译生成,他们也不太可能意识到误译的风险。

7 小结

随着机器翻译质量的提高,无论是用于直接交流还是作为翻译工作流程的一部分,这项技术都促进了更多交流。然而,机器翻译开发者、翻译客户、翻译机构、译者和翻译消费者都需要考虑道德伦理问题。与所有技术一样,机器翻译开发和输出都不应被视为道德中立,而是传播开发人员或创造翻译数据的译者的观点,他们的观点体现在为与机器翻译交互而设计的工具以及机器译文之中。对评估结果良好的机器翻译不加批判的报道,淡化了公众对机器译文存在风险和偏见的意识,同时有可能贬低人工翻译。在使用和处理机器翻译时,读者不妨参考本章提出的问题,并思考相关过程是否符合自己的价值观、目标和原则。

本章先从翻译来源和译者数据、所有权、许可、版权和流通模式等方面来讨论道德问题。在翻译工作流程中,机器翻译的使用是否合乎道德规范可能取决于所有利益相关者的态度、保密规则以及机器翻译平台背后的设计决策。这与可持续性、人机互动模式以及译者对工作流程的自主掌握程度有关。我们对于能够确保多样性和消除偏见的系统和数据的方法或许最为欠缺,但毫无疑问,这需要进一步的讨论和调整。

References

- Abdallah, Kristiina. 2010. Translator's agency in production networks. In Tuija Kinnunen & Kaisa Koskinen (eds.), *Translator's agency*, 11–46. Tampere: Tampere University Press.
- Baker, Mona. 2018. *In other words*. London: Routledge.
- Canfora, Carmen & Angelika Ottmann. 2020. Risks in neural machine translation. *Translation Spaces* 9(1). 58–77.
- Chesterman, Andrew. 2001. Proposal for a hieronymic oath. *The Translator* 7(2). 139–154. DOI: 10.6509.2001.10799097.

- Cronin, Michael. 2017. *Eco-translation: Translation and ecology in the age of the anthropocene*. London: Routledge.
- Docherty, Peter, Mari Kira & A. B. (Rami) Shari. 2008. What the world needs now is sustainable work systems. In Peter Docherty, Mari Kira & A. B. (Rami) Shari (eds.), *Creating sustainable work systems: Developing social sustainability*, 1–22. London: Routledge.
- European Parliament. 1996. *Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases*. <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:31996L0009>.
- Floridi, Luciano & Mariarosaria Taddeo. 2016. *What is data ethics?* DOI: 10.1098/rsta.2016.0360.
- Freitag, Markus, George Foster, David Grangier, Viresh Ratnakar, Qijun Tan & Wolfgang Macherey. 2021. Experts, errors, and context: A large-scale study of human evaluation for machine translation. *Transactions of the Association for Computational Linguistics* 9. 1460–1474. DOI: 10.1162/tacl_a_00437.
- ISO. 2015. *ISO 17100:2015. Translation services – requirements for translation services*. <https://www.iso.org/standard/59149.html>.
- ISO. 2017. *ISO 18857:2017. Translation services – post-editing of machine translation output: Requirements*. <https://www.iso.org/standard/62970.html>.
- Johnson, Marvin. 2020. *A scalable approach to reducing gender bias in Google Translate*. <https://ai.googleblog.com/2020/04/a-scalable-approach-to-reducing-gender.html> (20 May, 2020).
- Kay, Martin. 1980. *The proper place of men and machines in language translation* (Report CSL- 80-11). Palo Alto, CA: Xerox Corporation.
- Lambert, Joseph. 2018. How ethical are codes of ethics? Using illusions of neutrality to sell translations. *Journal of Specialised Translation* 30. 269–290. https://www.jostrans.org/issue30/art_lambert.php.
- Moorkens, Joss. 2020. “A tiny cog in a large machine”: Digital Taylorism in the translation industry. *Translation Spaces* 9(1). 12–34.
- Moorkens, Joss & David Lewis. 2019a. Copyright and the reuse of translation as data. In Minako O’Hagan (ed.), *The Routledge handbook of translation and technology*, 469–481. London: Routledge.
- Moorkens, Joss & David Lewis. 2019b. Research questions and a proposal for the future governance of translation data. *Journal of Specialised Translation* 32. 2–25.
- Moorkens, Joss & Martha Rocchi. 2021. Ethics in the translation industry. In Kaisa Koskinen & Nike K. Pokorn (eds.), *The Routledge handbook of translation and ethics*, 320–337. London: Routledge.

- Pérez-Ortiz, Juan Antonio, Mikel L. Forcada & Felipe Sánchez-Martínez. 2025. 神经机器翻译的工作原理. In Dorothy Kenny (ed.), 机器翻译知识普及: 为人工智能时代的用户赋能, 111–128. Berlin: Language Science Press. DOI: 10.5281/zenodo.14922297.
- Pym, Anthony. 2012. *On translator ethics: Principles for mediation between cultures*. Amsterdam: John Benjamins.
- Rossi, Caroline & Alice Carré. 2025. 如何选择合适的神经机器翻译解决方案: 机器翻译质量评测. In Dorothy Kenny (ed.), 机器翻译知识普及: 为人工智能时代的用户赋能, 39–63. Berlin: Language Science Press. DOI: 10.5281/zenodo.14922289.
- Sakamoto, Akiko. 2019. Unintended consequences of translation technologies: From project managers' perspectives. *Perspectives* 27(1). 58–73.
- Steinberger, Ralf, Andreas Eisele, Szymon Kloczek, Spyridon Pilos & Patrick Schlüter. 2012. DGT-TM: A freely available translation memory in 22 languages. In *Proceedings of the 8th international conference on language resources and evaluation (LREC 2012)*. ELRA. <https://www.aclweb.org/anthology/L12-1481/>.
- Strubell, Emma, Ananda Ganesh & Andrew McCallum. 2019. *Energy and policy considerations for deep learning in NLP*. <https://aclanthology.org/P19-1355.pdf>.
- Tiedemann, Jörg. 2012. *Parallel data, tools and interfaces in OPUS*. In *Proceedings of the 8th International Conference on Language Resources & Evaluation (LREC 2012)*. 2214–2218. Luxembourg: ELRA. http://www.lrec-conf.org/proceedings/lrec2012/pdf/463_Paper.pdf.
- Tomalin, Marcus, Bill Byrne, Shauna Concannon, Danielle Saunders & Stefanie Ullmann. 2021. *The practical ethics of bias reduction in machine translation. Why domain adaptation is better than data debiasing*. 419–433. DOI: 10.1007/s10676-021-09583-1.
- Topping, Suzanne. 2000. Sharing translation database information: Considerations for developing an ethical and viable exchange of data. *Multilingual* 11(5). 59–61.
- Toral, Antonio. 2019. Post-edits: An exacerbated translationese. In *Proceedings of machine translation summit XVII*, 273–281. EAMT. <https://www.aclweb.org/anthology/W19-6627/>.
- Troussel, Jean-Christophe & Julien Debussche. 2014. *Translation and intellectual property rights*. Luxembourg: Publications Office of the European Union.
- Van Wynsberghe, Aimee. 2021. Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics* 1. 213–218. DOI: 10.1007/s43681-021-00043-6.
- Vanmassenhove, Eva. 2019. *On the integration of linguistic features into statistical and neural Machine translation*. PhD Thesis. Dublin City University.

- Vanmassenhove, Eva, Dimitar Shterionov & Andy Way. 2019. Lost in translation: Loss and decay of linguistic richness in machine translation. In *Proceedings of machine translation summit XVII: Research track*. Dublin: EAMT, 222–232. <https://www.aclweb.org/anthology/W19-6622.pdf>.
- Vieira, Lucas Nunes, Minako O'Hagan & Carol O'Sullivan. 2020. Understanding the societal impacts of machine translation: A critical review of the literature on medical and legal use cases. *Information, Communication & Society* 24(11). 1515–1532. DOI: 10.118X.2020.1776370.
- Wachter, Sandra & Brent Mittelstadt. 2019. A right to reasonable inferences: Rethinking data protection law in the age of big data and AI. *Columbia Business Law Review* 2019(2). 1–130.
- Weizenbaum, Joseph. 1986. Not without us. *Computers and Society* 16. 2–7.
- Whipple, Judith M., Daniel F. Lynch & Gilbert N. Nyaga. 2010. A buyer's perspective on collaborative versus transactional relationships. *Industrial Marketing Management* 39(3). 507–518. DOI: 10.1016/j.indmarman.2008.11.008.
- World Intellectual Property Organization. 1979. *Berne convention for the protection of literary and artistic works*. (as amended on September 18, 1979). Geneva: WIPO. <http://www.wipo.int/wipolex/en/details.jsp?id=12214>.