

TAUS 指南在快速译后编辑 (LPE) 中的应用 ——以医学报告为例

仲晨阳, 倪蓉

(上海理工大学 外语学院, 上海 200093)

摘要: 随着互联网、大数据、云计算的飞速发展, 机器翻译发生了巨大变化, 机器翻译译后编辑 (machine translation post-editing, MTPE) 模式也应运而生, 逐渐成为翻译服务业的主流模式。但若要进一步提高 MTPE 效率和质量, 还需要相应的规范和准则加持, 以帮助译后编辑者识别机器翻译错误并快速决策机器译文是否需要编辑。通过介绍翻译自动化用户协会 (TAUS) 的 MTPE 指南的相关实践准则, 并根据文本性质和受众选取适用准则指导进行了医学报告的快速译后编辑 (light post-editing, LPE) 实践, 对医学报告文本译后编辑实践中存在的词汇错译、分词短语错译、从句错译以及段落错译、漏译等问题进行了分析, 并对该指南的实际指导意义进行了总结与评价。

关键词: 机器翻译; 快速译后编辑; 机器翻译译后编辑指南; 医学报告翻译

中图分类号: H 059 **文献标志码:** A **文章编号:** 1009 - 895X(2022)03 - 0238 - 07

DOI: 10.13256/j.cnki.jusst.sse.2022.03.005

Applying MTPE Guidelines of TAUS in Light Post-editing (LPE) —A Case of a Medical Report

ZHONG Chenyang, NI Rong

(College of Foreign Languages, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: With rapid development of the Internet, big data, and cloud computing, machine translation has undergone tremendous changes, which gives birth to machine translation post-editing (MTPE) model and helps it to be the mainstream model in translation service industry. However, further improvement of the efficiency and quality of MTPE also needs implementation of corresponding specifications and guidelines, so as to help post-editors identify errors in machine translation and decide whether it needs to be edited. This article introduces the MTPE guidelines by TAUS (the Association of Translation Automation Users) which guides the light post-editing (LPE) practice of a medical report, analyzes existing problems during the post-editing process, and evaluates the significance of the guidelines.

Keywords: machine translation; light post-editing; MTPE guidelines; medical report translation

近年来,机器翻译质量得到了极大改善,信息技术已经与翻译服务业深度融合。同时传统的纯人工翻译无法满足全球化和本地化催生的海量翻译需求,且成本相对较高,新的翻译业务类型与质量标准相较以前都有了很多变化,从而促使越来越多的企业开始用机器翻译技术进行初步的翻译项目处理,也就导致了译后编辑市场的扩大。作为提升机器翻译质量的新模式,译后编辑还缺少规范一致的实践原则、质量评估标准等,因此若要该模式得到有效并广泛的应用,还需要清晰明确的各类指导准则。2016年,翻译自动化用户协会(Translation Automation User Society, TAUS)发布了MT Post-editing Guidelines,这是目前可参考的较为完整的机器翻译译后编辑指南。本文根据该指南对医学报告机器译文进行了快速译后编辑实践,并对其指导意义及存在的问题进行了分析。

一、机器翻译与译后编辑

(一) 机器翻译的发展及缺陷

机器翻译的广泛应用产生了巨大的社会效益和经济效益。回顾其发展历程,主要经历了四个阶段:一是基于规则(RBMT);二是基于例子(EBMT);三是统计机器翻译(SMT);四是神经网络机器翻译(NMT)。神经网络机器翻译利用已有的大规模的真实语料库来进行深度学习,从语料库中自动获取语言特征和规则,它是基于大数据、使用神经网络来实现翻译的机器翻译系统^[1]。神经网络机器翻译的发展使机器译文质量得到了质的提升。

但是,即便神经机器翻译系统使机器译文质量飞速发展,机器译文仍有很多局限性,自然语言处理仍有许多在实践中难以解决的问题,比如从句错译、词汇错译、译文调序失败、符号错译、漏译等问题^[2]。医学报告作为医学文本的一种,其特点就是缩略语、医学术语使用较多,同时为表达客观及行文简洁准确,被动语态以及分词结构应用较多^[3]。机器在翻译医学文本时,往往会出现术语错译、漏译和词汇替代错译问题,同时也会出现对被动句以及分词结构的处理不当导致译文含义与原文出现偏差的现象。具体可见本文第三节的实践分析。要克服机器译文存在的这些缺陷,使译文质量达到要求,目前最快捷有效的方法就是对其进行人工修改、完善,即进行机器翻译译后编辑。

(二) 译后编辑概述

译后编辑(post-editing)是在语言或格式方面,对机器翻译的原始产出,即初始译文,进行加工与修改来提高机译产出的准确性与可读性^[4]。2010年TAUS实践中的译后编辑报告将译后编辑定义为“用最少的人工改进机器生成的翻译的过程”^[5]。而针对不同要求和目的,关于机器翻译译后编辑的ISO 18 587标准将译后编辑分为两个级别:快速译后编辑(Light Post-editing, LPE)、完全译后编辑(Full Post-editing, FPE)^[6]。机器翻译译后编辑模式充分提高机器翻译的速度(效率),也保证发挥人工翻译的精度(质量),从而既满足翻译市场快速发展的需求,也推动了翻译技术的发展,还促进了学界和业界的交流与合作,丰富了语言服务产业链的组成^[7]。

不过译后编辑作为新的翻译工作模式,在实践过程中除了要识别并纠正上述机器译文各类错误之外,还需考虑效率、质量要求、成本等各方面因素。因此译后编辑在发展过程中还需要独立、一致的标准来衡量和约束译后编辑的质量,确定译后编辑工作量等,帮助译后编辑者更加高效高质地完成译后编辑任务,同时培养出更多高质量的译后编辑人员。但目前机器翻译译后编辑相关的大部分实践准则就比较宏观,如崔启亮提出的实践准则相对宏观,进行实践时还需自行确立具体规则^[7]。而Midori Tatsuni提出的对MTPE译文相关的部分要求与TAUS类似,但比较宽泛,并未针对不同的质量要求进行细致划分,实际应用时还需综合考虑各种因素进行细化^[8]。目前相对具体的实践准则是TAUS发布的MT POST-EDITING GUIDELINES,根据不同的质量要求列出了对应的原则,所以本文采用该原则指导所选文本的译后编辑实践,以检验其是否能有效提高译后编辑效率。

二、TAUS 机器翻译译后编辑指南

TAUS成立于2004年,是全球语言和翻译行业的资源中心。该协会通过自己的数据云和质量评估服务为翻译行业提供相关的建议、工具、指标、基准和数据等。2016年,该协会发行了MTPE指南,旨在促进译后编辑模式的发展,提升译后编辑的质量和效率,并帮助该行业选择、培训高素质、高水准的译后编辑工作者。

该指南指出,最基本的译后编辑质量评估准则有两条,一是机器生成的译文质量,二是对待翻译材料的最终质量预期,即译后编辑工作如何进行,取决于机器生成的原始译文质量及客户对译文质量

的需求^[9]。该指南把预期的最终译后编辑质量大致分为两个等级,一级为“good enough quality”,另一级为“human translation quality”,其具体要求见表1。

表1 译后编辑质量要求
Tab.1 Requirements for MT Post-editing

质量等级	Good Enough Quality	Human Translation Quality
质量要求	(1) 最终译文语义正确;	(1) 最终译文要保证语法、句法、语义正确;
	(2) 确保无信息被无故增加或遗漏;	(2) 确保除客户指定的“无需翻译”术语表中的术语,其他术语均翻译准确;
	(3) 对任何有害、不妥及文化上不可接受的信息进行编辑;	(3) 确保无信息被无故增加或遗漏;
	(4) 尽可能多地保留机器生成的原始译文;	(4) 对任何有害、不妥及文化上不可接受的信息进行编辑;
	(5) 确保最终译文遵守拼写相关的基本适用规则;	(5) 尽可能多地保留机器生成的原始译文;
	(6) 无需进行仅跟文风相关的改正;	(6) 确保最终译文遵守拼写、标点符号、连字符相关的基本适用规则;
	(7) 不需仅为改善译文流畅度进行句子重组。	(7) 确保最终译文格式正确。

在选择适用质量要求时,主要取决于目标读者或客户对最终译文质量的需求(包括最终译文的受众或使用目的)以及初始机器译文的质量。而对于初始机器译文质量的评判,若花费太多时间来确定其是否可用会得不偿失,降低效率。因此可在决策时间上加一些限制,若在一定时间内查看一个机器翻译片段(在熟悉源文本和目标文本之后),发现无法轻松理解,那就舍弃机器译文。Mesa-Lao在其实验分析中表明,大部分测试者在初读机器译文上的停留时间为5~10秒左右,本文的译后编辑实践即以5~10秒原则来判断是应该纠正机器译文,还是应该删除并重新翻译低质量的片段^[10]。同时针对最终译文质量的不同质量预期(是“good enough quality”还是“human translation quality”)以及机器生成译文的不同质量,采取不同的译后编辑策略,决定是进行LPE还是FPE。

三、基于TAUS指南进行的医学文本译后编辑实践分析

医学报告属于医学类文本,作为科技文本的一种,对MTPE模式的适应性较强,同时有海量翻译需求,对翻译效率要求较高。本文选取了*The New England Journal of Medicine*上的医学报告来进行译后编辑实践,这是关于2019新型冠状病毒肺炎武汉病例分析报告,发布于2020年1月24日,全文约

3000词。此文的目的是与医学工作者和研究者共享信息以及及时有效共同应对并预防疫情。因为该流行病传染性强,传播速度快,加之如今交通非常发达,人员流动频繁,所以此类文本信息翻译非常注重准确性和时效性。同时,此类文本译文目标受众为医学专家或工作者,他们仅需了解原文概述或要点,以帮助自己了解疾病相关信息,共同做好应对准备并寻找解决办法,所以可选择省时高效的LPE模式进行MTPE,使最终译文达到“good enough quality”的要求(下文提到的具体准则均针对“good enough quality”)。

本文所选择的机器翻译引擎为谷歌翻译,在实践过程中发现词汇错译、分词短语错译、从句错译以及段落错译、漏译问题较多,而此前一直被视为难点的被动语态问题已不明显。基于神经网络机器翻译强大的学习能力以及大数据的支撑,目前谷歌对被动句的处理表现较为良好,比如,笔者较早时期实践时发现的相关问题,在几个月之后再次用谷歌生成译文时,发现其已经对译文进行了改良,符合“good enough quality”的要求,可不再作为难点进行分析。其他相关案例具体分析如下。

(一) 词汇错译

例1 In late December 2019, several local health facilities reported clusters of patients with pneumonia of unknown cause that were epidemiologically linked to a seafood and wet animal wholesale market in

Wuhan, Hubei Province, China. ...We report the results of this investigation, identifying the source of the *pneumonia clusters*, and describe a novel coronavirus detected in patients with pneumonia whose specimens were tested by the China CDC at an early stage of the outbreak.

谷歌译文 (2020年8月9日) 2019年12月下旬,几家地方卫生机构报告了一群原因不明的肺炎,其在流行病学上与中国湖北省武汉市的海鲜和湿动物批发市场有关。……我们报告了这项调查的结果,确定了肺炎簇的来源,并描述了在疾病爆发早期由中国疾病预防控制中心检测其标本的肺炎患者中检测到的新型冠状病毒。

修改译文 2019年12月下旬,中国湖北武汉几家当地的医疗机构报告了一群原因不明的肺炎患者,从流行病学角度看,均与当地的一家海鲜和水产动物批发市场有关。……我们报告了本次的调查结果,确定了肺炎患者群体的来源,并描述了在疾病爆发早期由中国疾病预防控制中心检测其标本的肺炎患者中检测到的新型冠状病毒。

通过5~10秒原则可快速识别出谷歌关于“wet animals”和“clusters”的翻译出现了错误,导致句意出现偏差,为达到要求(1)“语义正确”,对这两处进行了编辑,而“clusters”也按照上下文语境选择正确的词义。同时对比原文还发现,谷歌译文出现了信息遗漏的现象,漏译了“clusters of patients”,因此为达到TAUS指南要求(2)“确保信息无遗漏”,确保译文信息无添加或遗漏,在进行快速译后编辑时将遗漏信息补全。

例2 *Emerging and reemerging pathogens* are global challenges for public health. Coronaviruses are enveloped RNA viruses that are distributed broadly among humans, other mammals, and birds and that cause respiratory, enteric, hepatic, and neurologic diseases.

谷歌译文 (2020年8月9日) 病原体的出现和重新出现是公共卫生的全球性挑战。冠状病毒是被包膜的RNA病毒,广泛分布于人类,其他哺乳动物和鸟类中,并引起呼吸系统,肠道,肝脏和神经系统疾病。

修改译文 新兴和再现病原体是公共卫生的全球性挑战。冠状病毒是有包膜RNA病毒,广泛分布于人类、其他哺乳动物以及鸟类宿主中,并可引起呼吸道、肠道、肝脏和神经系统疾病。

结合5~10秒原则以及医学文本翻译的重难点,可观察到“病原体的出现和重新出现”这里是不够准确的。“emerging”有新兴的、出现的、形成的等意思,但在本例中该词应取“新兴的”一意,谷歌在选择词义时出现错误;因此为符合TAUS指南第(1)条要求,对语义不正确的地方进行快速译后编辑,其他部分机器译文基本符合TAUS指南的质量标准,因此按第(4)条要求保留了大部分原始机器译文。

例3 Human airway epithelial cell cultures were generated in an air-liquid interface for 4 to 6 weeks to form well-differentiated, polarized cultures resembling in vivo pseudostratified mucociliary epithelium.

谷歌译文 (2020年8月9日) 人气道上皮细胞培养物在气液界面上产生4至6周,形成分化良好的极化培养物,类似于体内假复层粘膜纤毛上皮细胞。

修改译文 在气液相界面培养人类呼吸道上皮细胞4至6周,形成分化良好的极化培养物,类似于体内假复层粘膜纤毛上皮细胞。

借助5~10秒原则,本例中的“generate”有“繁殖、产生、发生”等意思,在本句中应取“繁殖”的含义。谷歌在选择词义时未能准确根据语境识别,使译文句意产生偏差不易理解,因此在按照TAUS指南要求进行快速译后编辑时对该词的翻译进行编辑。而此处指的是“人类呼吸道上皮细胞被繁殖”,因此在译后编辑时根据句意将“generate”译为“培养”。

(二) 分词短语错译

例4 More than 20,000 viral reads from individual specimens were obtained, and most contigs matched to the genome from lineage B of the genus betacoronavirus — showing more than 85% identity with a bat SARS-like CoV (bat-SL-CoVZC45, MG772933.1) genome published previously.

谷歌译文 (2020年8月9日) 从单个标本中获得了20 000多个病毒读物,并且大多数重叠群与beta冠状病毒属B谱系的基因组相匹配——与蝙蝠SARS样冠状病毒(bat-SL-CoVZC45, MG772933.1)的同源性超过85%基因组先前已发表。

修改译文 从单个标本中获得了20 000多个病毒读段,大多数重叠群与β冠状病毒属B谱系的基

基因组相匹配——与之前发布的蝙蝠 SARS 样冠状病毒 (bat-SL-CoVZC45, MG772933.1) 基因组的同源性超过 85%。

本例中的“published previously”作为 bat SARS-like CoV genome 的后置定语,谷歌译文未对语序进行调整,保留了其原文的序位,导致译文产生歧义,不够准确,所以为符合 TAUS 指南要求(1),对其进行了调整。同时,“reads”一词既有普通释义又有专业释义,需紧扣上下文关系,选择医学释义^[11]。借助 5~10 秒原则也可快速判断 reads 在本篇医学报告中译为“读物”是不正确的,所以为达到 TAUS 指南要求,对其进行修改。通过术语在线查询该词,发现其在生物物理学领域是“测序片段”“测序读段”的意思,因此将其译为“读段”。

例 5 Our study showed that initial propagation of human respiratory secretions onto human airway epithelial cell cultures, followed by transmission electron microscopy and whole genome sequencing of culture supernatant, was successfully used for visualization and detection of new human coronavirus that can possibly elude identification by traditional approaches.

谷歌译文 (2020年8月9日) 我们的研究表明,人呼吸道分泌物在人气道上皮细胞培养物中的初步繁殖,然后通过透射电子显微镜和培养物上清液的全基因组测序,已成功用于可视化和检测新的人冠状病毒,而这可能无法通过传统方法进行鉴定。

修改译文 我们的研究表明,在人类呼吸道细胞培养基中对人类呼吸道分泌物进行初步繁殖,然后用透射电子显微镜观察并对培养物上清液进行全基因组测序,可成功可视化并检测新型人类冠状病毒,而这种病毒可能无法通过传统方法进行鉴定。

本例中的分词短语作插入语,谷歌对该插入语的处理按照原文顺序和词性生成了机器译文,但逻辑上不够连贯造成语义不够准确,引起了理解上的困难,无法达到 TAUS 指南要求(1)。而 Cook 曾提出要保证从读者的认知角度看,译文语篇要连贯以免给读者的信息提取设置障碍和干扰^[12]。因此在译后编辑时根据前后语境补充了语义,并对整个句子进行了逻辑重组。此外,本例还存在术语不一致、词义选择错误以及代词指代不清等问题,不符合 TAUS 相关指南要求的质量,因此均对其进行了编辑。

(三) 从句错译

例 6 Four lower respiratory tract samples, in-

cluding bronchoalveolar-lavage fluid, were collected from patients with pneumonia of unknown cause who were identified in Wuhan on December 21, 2019, or later and who had been present at the Huanan Seafood Market close to the time of their clinical presentation.

谷歌译文 (2020年8月9日) 从2019年12月21日或以以后在武汉发现的,原因不明的肺炎患者中收集了四个下呼吸道样本,包括支气管肺泡灌洗液,这些样本在他们离开时已出现在华南海鲜市场临床表现。

修改译文 我们从2019年12月21日前后在武汉确诊的不明原因肺炎患者身上收集了四份下呼吸道样本(包括支气管肺泡灌洗液),他们出现临床症状不久前都曾去过华南海鲜市场。

借助 5~10 秒原则,可识别出本句最后一句的机器译文是无法理解的,回看原文可发现本句有两个定语从句均修饰“patients”,谷歌在处理时并未准确识别本句逻辑,将第二个从句判定为是修饰句首“samples”的,导致最终生成的译文句意产生偏差,不符合 TAUS 指南的要求(1)。因此在译后编辑时对其进行了修改,将第二个从句的译文予以修正,并为达到要求(4)对第一个从句的译文进行了 LPE 处理。

例 7 Our study showed that initial propagation of human respiratory secretions onto human airway epithelial cell cultures, followed by transmission electron microscopy and whole genome sequencing of culture supernatant, was successfully used for visualization and detection of new human coronavirus that can possibly elude identification by traditional approaches.

谷歌译文 (2020年8月9日) 我们的研究表明,人呼吸道分泌物在人气道上皮细胞培养物中的初步繁殖,然后通过透射电子显微镜和培养物上清液的全基因组测序,已成功用于可视化和检测新的人冠状病毒,这可能无法通过传统方法进行鉴定。

修改译文 我们的研究表明,在人类呼吸道细胞培养基中对人类呼吸道分泌物进行初步繁殖,然后用透射电子显微镜观察并对培养物上清液进行全基因组测序,可成功可视化并检测新型人类冠状病毒,而这种病毒可能无法通过传统方法进行鉴定。

本句较长,借助 5~10 秒原则可发现最后一句语义不清晰,易产生歧义。仔细阅读原文,发现谷歌在处理修饰“new human coronavirus”的定语从句时,没有明确引导词的指代,使译文句意模糊不

够准确,为达到TAUS指南要求(1),对该处进行了编辑。同时机器译文还出现了词性未能准确转换导致译文意思不够明确的现象,即对“initial propagation”的处理不够灵活准确,所以译后编辑时将其修改为“进行初步繁殖”。

(四)段落错译、漏译问题

例8 Three adult patients presented with severe pneumonia and were admitted to a hospital in Wuhan on December 27, 2019. Patient 1 was a 49-year-old woman, Patient 2 was a 61-year-old man, and Patient 3 was a 32-year-old man. Clinical profiles were available for Patients 1 and 2. Patient 1 reported having no underlying chronic medical conditions but reported fever (temperature, 37 °C to 38 °C) and cough with chest discomfort on December 23, 2019. *Four days after the onset of illness*, her cough and chest discomfort worsened, but the fever was reduced; a diagnosis of pneumonia was based on computed tomographic (CT) scan. ...Patients 1 and 3 recovered and were discharged from the hospital on January 16, 2020. Patient 2 died on January 9, 2020. No biopsy specimens were obtained.

谷歌译文(2020年10月7日) 三名患有严重肺炎的成年患者于2019年12月27日入武汉医院。患者1为49岁女性,患者2为61岁男性,患者3为32岁老人患者1和2可获得临床资料。患者1报告于2019年12月23日无基本慢性病,但报告有发烧(温度37 °C至38 °C)和咳嗽伴胸口不适。疾病,咳嗽和胸部不适加剧,但发烧减少;肺炎的诊断基于计算机断层扫描(CT)扫描。……患者1和3已康复,并于2020年1月16日出院。患者2于2020年1月9日死亡。未获得活检标本。

修改译文 三名患有严重肺炎的成年患者于2019年12月27日入武汉医院。患者1为49岁女性,患者2为61岁男性,患者3为32岁男性。患者1和2可获得临床资料。患者1报告于2019年12月23日,无基本慢性病,但报告有发烧(温度37 °C至38 °C)和咳嗽伴胸口不适。患者1在发病四天后,咳嗽和胸部不适加剧,但发烧减少,肺炎的诊断基于计算机断层扫描(CT)扫描。……患者1和3已康复,并于2020年1月16日出院。患者2于2020年1月9日死亡,未获得活检标本。

根据5~10秒评价原则,可以快速判定“Pa-

tient 3 was a 32-year-old man”的译文是有问题的,然后对其进行了修改。但令人奇怪的是前两处对于年龄的翻译处理都很准确,仅此处译文出现了问题,为进行确认,笔者使用谷歌在线翻译引擎尝试了3~4次,都得到了同样的结果。而本段中“Four days after the onset of illness”被遗漏了,这里的漏译仅凭5~10秒阅读机器生成的译文是判定不出来的,必须结合原文进行审校检查。但其实该句语法并不复杂,鉴于之前谷歌译文的良好表现,笔者为确认也进行了3~4次的谷歌翻译尝试,得出的结果均相同。鉴于此,笔者认为这是阻碍译后编辑效率提高的一个重要因素,也是机器翻译应尽快解决的问题。

综上,在本次实践过程中发现,相当一部分机器原始译文已符合TAUS指南中“good enough quality”的要求,不用做太多修改。且机器翻译质量在持续提高,在医学专有名词、术语以及被动语态方面均表现良好,但对一词多义的词汇识别还不够准确,问题较多,同时对长句、结构复杂的句子处理较为糟糕,如定语从句、分词结构等,转换时无法对句子语序进行调整,导致译文意思出现偏差,甚至在翻译段落时出现漏译现象。

通过使用TAUS指南指导LPE的实践发现,指导性较强的是要求(1)(2)(4),可根据此三条要求快速做出判断是否应对机器译文进行LPE。因本次实践中未遇到文化上不妥和侵犯信息,以及因本文为英译汉,汉语无拼写问题,所以要求(3)和(5)关于文化和拼写问题不用讨论。要求(4)在应用时应注意尽可能多地保留机器初始译文是建立在机器译文质量较高的基础上,译者应迅速做出判断是否舍弃机器译文进行人工翻译,否则时间成本反而更高。要求(6)无需进行仅跟文风相关的改正和(7)不需仅为改善译文流畅度进行句子重组,就允许我们保留一些不符合中文语言习惯但句子含义表达正确的机器译文而无需进行修改。如例1仅修改了词汇错译的部分,其余均保留机器译文,节约了很多时间。通过使用TAUS指南,在进行译后编辑时会更具有目的性,标准更加明确,可尽量避免做一些不必要的修改导致时间成本和人工成本的浪费。当然,TAUS指南仍有许多不足之处,依旧不够具体,有待在实践应用中继续完善。

四、结束语

语言服务行业采用机器翻译译后编辑模式来提

高翻译效率和翻译质量,也需要统一规范的标准来帮助评估工作质量并定价。同时好的规范还可以促进 MTPE 模式效率和质量的进一步提升,如本文按照 TAUS 指南对“good enough quality”的要求进行快速译后编辑时,目标更加明确和清晰,避免译后编辑过程中因标准不确定进行不必要的编辑,导致比纯人工翻译耗时更长的情况发生。且此类医学报告信息交流的时效性尤为关键,因为流行病的传播快、传染性强,信息能够更快实现共享,就可能更早更有效地防控疾病更大范围传播,就可能帮助各国医务工作者共同合作更早研究出应对方法。因此,采取较高的机器初始译文质量加清晰的译后编辑要求(规则),可帮助译后编辑者们在更短时间内获得符合预期的译文。当然本文所遵循的 TAUS 指南也依旧存在需要改善或进一步细化的地方,相信随着机器翻译译后编辑的发展,更加完善的规范会生成,机器翻译译后编辑模式也会更加成熟高效。

参考文献:

- [1] 冯志伟. 机器翻译与人工智能的平行发展[J]. 外国语, 2018, 41(6): 35-48.
- [2] 罗季美, 李梅. 机器翻译译文错误分析[J]. 中国翻译, 2012, 33(5): 84-89.
- [3] 刘玉强, 贺子琼. 医学英语的文体特征及其翻译策略[J]. 英语广场, 2015(10): 35-36.
- [4] 冯全功, 崔启亮. 译后编辑研究: 焦点透析与发展趋势[J]. 上海翻译, 2016(6): 67-74, 89.
- [5] TAUS. Post-editing in Practice[EB/OL]. (2010-03-10) [2020-08-20]. <https://www.taus.net/think-tank/reports/postedit-reports/postediting-in-practice>.
- [6] ISO. ISO 18587: 2017 Translation services — Post-editing of machine translation output — Requirements[S]. ISO, 2017.
- [7] 崔启亮. 论机器翻译的译后编辑[J]. 中国翻译, 2014, 35(6): 68-73.
- [8] TATSUMI M. Post-editing machine translated text in a commercial setting: observation and statistical analysis[D]. Ireland: Dublin City University, 2010.
- [9] TAUS. TAUS MT Post-editing guidelines[EB/OL]. (2016-01-20)[2020-08-20]. <https://www.taus.net/think-tank/reports/postedit-reports/taus-post-editing-guidelines>.
- [10] SANCHIS-TRILLES G, ALABAU V, BUCK C, et al. Interactive translation prediction versus conventional post-editing in practice: a study with the CasMaCat workbench[J]. Machine Translation, 2014, 28(3): 217-235.
- [11] 李拜华. 英汉医学翻译中的词义选择[J]. 中国翻译, 1995(1): 16-22.
- [12] COOK G. Discourse and Literature[M]. Shanghai: Shanghai Foreign Language Education Press, 1999.

(编辑: 朱渭波)