

公示材料

一、基本信息			
项目名称	中文	专业标准化中英双语智能翻译系统构建研究及应用	
	英文	Research and application of professional Chinese-English bilingual intelligent translation system for standardization	
成果申报等级		<input checked="" type="checkbox"/> 一等奖 <input type="checkbox"/> 二等奖 <input type="checkbox"/> 三等奖	<input checked="" type="checkbox"/> 是否同意调级
主要完成人		刘智洋、张宝林、甘克勤、蒙永业、朱励、田方、崔启亮、邵珊珊、陈劲舟、高燕	
主要完成单位		中国标准化研究院、国家市场监督管理总局国家标准技术审评中心、北京悦尔信息技术有限公司	
推荐单位(盖章)		中国标准化研究院	
奖项的主要项目来源		<input checked="" type="checkbox"/> 国家级 <input type="checkbox"/> 省部级 <input type="checkbox"/> 其他	
具体计划、基金的名称和编号：国家重点研发计划 NQI 专项课题“标准化英汉双语语料库及本体知识系统建设” 课题编号：2016YFF0202905			
成果的主要项目起止时间		起始： 2016-7	完成：2019-12
组织验收/鉴定单位		中国标准化研究院	
成果登记号	G2021-125	成果登记时间	2021 年 5 月 12 日

二、奖项简介

（主要技术内容、技术指标、创新点、授权知识产权情况、应用推广及取得的经济、社会效益等；限 1 页）

标准是世界的通用语言，是联通世界的桥梁纽带。《国家标准化发展纲要》明确要求“标准化国际合作深入拓展”。要实现这一战略目标，语言是关键、也是瓶颈。国际标准化工作对语言的要求不仅是要达意，更要求准确，而且需要符合国际标准化工作的语义规律，这就对参与国家标准化的专家的英语语言能力提出了极高的专业化要求，本项目开创性地建立并解决了中国专家参与国际标准化活动语言难题的有效途径。“标准化中英双语智能翻译云平台”是在 NQI 专项“标准化英汉双语语料库及本体知识系统建设”课题支持下完成的，2019 年正式上线运行，为标准文本和标准化文献的中英文互译提供了实时在线、准确快捷、高效权威的翻译工具，为中国标准国际化打造了便利的基础设施。

项目构建了超过 6000 万中文字符、2200 万英文单词的英汉双语标准化专业语料库，充分涵盖标准化专业用词用语和各领域专业用语。基于标准化英汉双语语料库，采用神经网络机器翻译技术搭建了国内首个标准化中英双语智能翻译云平台；平台支持 WORD/EXCEL/PPT/PDF/TXT/HTM 等多种格式文本的中英文“格式化”互译，具有在线编辑和专用语料库建库等功能。平台运行稳定、功能完备，经绩效评价专家组评审，一致认为，具有国际先进水平。

项目主要创新和贡献如下：

- 1、首次提出基于 ISO/IEC/ITU 等国际标准化组织国际标准文本和“等同采用”的国家标准、行业标准文本，首创“全句——短语——词语——缩语——专用代码”的多层智能结构构建技术，采用 Coh-Metrix、LancsBox 等语料库语言特征分析工具研究标准语言特征，成功建立了我国第一个权威的标准化中英双语语料库和中英双语标准本体知识系统；
- 2、基于标准文本和标准化文献的标准结构化特点，提出了基于标准化多层结构专业语料库的语义关联智能驱动匹配技术，创新了标准化领域动态线性聚合神经网络机器翻译建模技术，采用基于动态线性聚合的 Transformer 模型，解决了现有技术中神经网络机器翻译 Transformer 模型在训练深层网络时底层网络难以实现充分训练、导致无法有效地训练更深层的网络结构的问题，有效解提升了标准文本和标准文献的翻译性能。
- 3、采用页面网格分区技术，自动识别“图、文、表、公式”等页面要素，研制成功文字译后内容自动匹配的全文“格式化”自动翻译、自适应排版技术，成功实现了文本、文献页面完美对应的双语互译。首次自主研制平行语料加工处理的国家标准 GB/T40035—2021《双语平行语料加工服务基本要求》，等同采用国际标准研制机器翻译国家标准 GB/T40036—2021《翻译服务 机器翻译的译后编辑 要求》，并于 2021 年 4 月获得国家标准委批准发布。

该成果获 2 项软件著作权。基于翻译并批准出版外文版标准超过 50 项。至今平台活跃的注册用户超过 5000 个，已完成近 2 亿中文文字或英文词汇互译。为抗击疫情支持复工复产免费提供近 1 亿字/词服务。2021 年开始试水收费服务，在各平台普遍提供免费服务的环境下获得市场认同，已实现近 80 万元的翻译直接收费，目前收费用户正在快速发展，为专业化运维提供了费用保障，实现了自身造血功能、以平台养平台的良性循环模式和长期可持续发展。