

建立语言能力标准化服务体系

——创新大数据平台建设模式

“一带一路”沿线 60 多个国家和民族讲着 20 多种不同语言，文化多元且差异巨大。随着“一带一路”国家战略的深入实施与推进，加强语言互通和信息相通，提升我国语言能力势在必行。语言能力标准化服务体系及大数据平台建设创新模式，通过将政、产、学、研各方联合在一起，助力打造“一带一路”国家战略研究的高地，满足国家和公民对于“一带一路”沿线国家和地区在克服相关语言障碍和大数据信息方面的需求，为促进我国广大企业“走出去”、加强创新能力及开放合作打造坚实的基础，为提升中国文化在“一带一路”国家认同感和感知力提供强劲动力。

一、建立语言能力标准化服务体系的背景

2016 年 5 月 15 日，陕西省人民政府外事办公室、西安外国语大学、中译语通科技股份有限公司和西咸新区秦汉新城管委会在 2016 丝绸之路博览会暨第 20 届东西部贸易洽谈会上签署《陕西省“一带一路”语言服务及大数据平台合作框架协议》，正式开启陕西省“一带一路”语言服务及大数据平台建设。通过构建语言能力标准化服务体系，建设“一带一路”大数据平台，为自贸试验区和陕西省与“一带一路”沿线国家的交流合作提供语言和信息支撑。

二、建立语言能力标准化服务体系主要做法及成效

（一）主要做法

1. 构建语言能力标准化服务体系

语言能力标准化服务体系包括语言服务规范、全球多语言呼叫中心及涉外服务人员外语培训三部分。

（1）确立语言服务规范标准

陕西省自贸试验区秦汉新城功能区引入中译语通“中国核心语汇”项目，对旅游景区、市政道路及机场酒店等公共场所的中文标志、说明进行整理及翻译工作，对已有公共场所外语标志进行专项整治，解决外语译写各自为政、一名多译、互相矛盾的问题，为城市规划、交通、旅游、市政、城建等部门和酒店、餐饮、商贸等各类企事业单位制作公示语和外文标志提供了统一规范。

（2）建立全球多语言呼叫中心

2017年5月15日，位于西咸新区秦汉新城的陕西省“一带一路”语言服务及大数据平台呼叫中心启动试运行。多语言呼叫中心基础知识库涵盖政治、经济、文化、旅游、医疗等领域，实现与政务、工商、旅游、招商引资、民政、医疗等涉外公共服务热线的对接，同时设立外语服务专线，具有多语种服务能力。全天候24小时运营，实现50路并发接入，响应时间不超过24秒。

（3）加强服务人员外语培训

中译语通选派资深翻译，定期对秦汉新城功能区交通、旅游、商业等重点领域服务行业从业人员开展培训讲座。举办涉外文明礼仪学习活动，把讲外语与讲文明、讲礼仪有机结合、相互促进，使市民的整体素质不断提高。

2. 建设“一带一路”大数据平台



陕西省“一带一路”语言服务及大数据平台

（1）平台建设情况

平台主要包括三大子系统，分别是跨语言大数据采集子系统、跨语言大数据处理子系统以及跨语言大数据分析应用子系统。通过跨语言大数据采集子系统将全球旅游舆情抓取过来，再经过跨语言大数据处理子系统的处理后，就可以利用跨语言大数据分析应用子系统进行相关数据的分析应用。

平台在“一带一路”沿线包括俄罗斯、印度、泰国、捷克、波兰等10多个

国家,及全球22个主要国家部署了服务器资源,全球共设置了32万个数据源挖掘点,涵盖200多个国家和地区,可以对全球数据源进行实时挖掘,日增3000万条数据。跨语言大数据挖掘子系统支持分布并发处理的能力,确保全球每个数据源数据更新的实时性,挖掘效率为200万篇文本/分钟。目前服务器已存储了100多亿网页数据,数据存储量100PB,可以扩充至日增长量10TB。

(2) 平台主要功能

海外智库:通过将海外尤其是“一带一路”沿线国家的相关数据分类、处理、关联后,形成了海外智库,可为政府和企业提供海外各国(城市)相关领域竞争力查询、海外各领域资讯查询、海外相关文献查询等功能。系统内嵌的基于神经网络的机器翻译引擎会自动将多语种信息翻译成中文,便于用户对感兴趣的数据信息进行中英文对照查看。

海外舆情实时监测与分析:平台自动、实时监测全球各领域舆情信息,并根据信息的热度自动进行热度舆情提醒显示,显示的舆情信息都由机器翻译引擎自动翻译成中文,便于用户对感兴趣的热点信息进行双语对照查看。政府也可根据舆情热点内容和传播速度快速反应、安全维稳。另外,用户还可针对某一海外区域、某一海外国家甚至某一海外城市,进行特定专题的舆情多维度分析,分析结果支持下载等多种转存方式。

大数据可视化展示:系统可以将各类数据进行分析建模,并利用可视化技术将海外旅游舆情大数据以图形化的形式展示出来。用户可以将其投放在展厅大屏,以便将数据醒目的展现和监控。

(二) 主要成效

“一带一路”语言服务标准体系及大数据平台建立以来,面向政府和企事业单位、外向型企业、个人提供了有效的服务,显示出巨大的社会和经济效益。

1. 助力打造国家语言战略智库

“一带一路”语言服务标准体系及大数据平台为陕西乃至全国培养了一批掌握“关键语言”的外语人才,建立起“语言互联网”,大大提高了民众个人语言能力(母语能力、多语能力)、社会语言能力(各种职业、专业语言能力)乃至国家语言能力(行使国家力量时所需的语言能力,如在抢险救灾、反恐维稳、海外维和、远洋护航、联合军演、护侨撤侨及各种国际合作中);为促进陕西自贸试验区企业“走出去”打下了坚实的基础,助力实施国家语言战略,更好地在“一带一路”沿线国家推广汉语和中国文化,创建中国周边语言文化数据库,打造国家语言战略智库。

2. 为企业“走出去”“引进来”提供便利

“一带一路”语言服务标准及大数据平台为陕西自贸试验区企业“走出去”提供基础的多语言服务保障,同时通过精准的大数据平台为企业的对外投资经营活动提供更为专业全面的数据引导,真正消除企业海外投资的各项信息壁垒。此外,平台作为信息和跨语言对接的重要窗口,促进国际企业了解陕西自贸试验区投资环境和特点,并对外商投资提供跨语言精准服务。

3. 为自贸试验区科学决策提供数据支撑

通过清洗挖掘和机器翻译功能,对收集到的数据进行加工处理,自贸试验区各级管理机构可以及时了解相关国家各领域运行情况。通过对获取数据深入分析挖掘、科学建模,逐步建立了各领域量化指标体系,定性定量评估各领域运行现状,分析推导数据间的关联,厘清发展脉络,为政府行政管理创新实践提供数据基础。通过抓取境内外社交媒体、搜索引擎及公众平台上各国民众及时、真实、原创的评论,能较为客观地映射各国民众心目中对自贸试验区各领域的感观印象,便于舆情监控。

三、建立语言能力标准化服务体系政策启示

语言能力标准化服务体系及大数据平台建设创新模式是自贸试验区西咸新区秦汉新城功能区率先推行的案例,已在全省进行复制推广,属于全国首创型创新举措。

(一) 构建完善的语言服务体系是促进国际人文交流的基础。语言服务能力是地区国际化发展的必备因素,也是促进与“一带一路”沿线国家人文交流合作的先决条件。语言标准化服务体系、多语呼叫中心等举措是提升地区语言服务水平的重要途径,对促进国际文化交流、带动地区国际旅游等方面起到重要作用。

(二) 建设中国特色的多语言交流平台可以为地区国际化推广提供有效助力。具有极强中国特色的“中国核心语汇”等项目,定制化省市专有语汇,可以有效提高地区的国际宣传力和影响力,是未来促进地区乃至全国进行国际文化交流的重要载体。

四、建立语言能力标准化服务体系下一步工作思路

(一) 进一步完善语言能力标准化服务体系

在持续推广“中国核心语汇”项目的同时,继续对其进行补充和完善,

制定陕西专属地方特色词汇，并按阶段、分批次补充“一带一路”沿线国家语言。

（二）帮助优秀文化作品“走出去”

优秀书籍作为文化输出的重要组成部分，对外国人了解中国历史、文化有着重要作用，陕西自贸试验区将联合陕西省内各大出版社，对国内优秀作品进行汇总，利用语言能力标准化服务体系，翻译成“一带一路”相关国家语种，在“一带一路”国家出版发行，让更多国家的人民了解中国。

（三）加强“一带一路”大数据平台建设

在“一带一路”沿线国家设置更多的服务器及布置更多的挖掘点，加大对机器翻译的功能提升，通过专业的跨语言大数据分析处理技术，完成舆情信息的智能语义分析和数据可视化展示，建立海外智库，为政府、企业制定相应决策提供支持和帮助。

【实践者说】

中译语通科技（陕西）有限公司常务副总经理王崇正说：“语言能力标准化服务体系及大数据平台建设创新模式，是中译语通科技（陕西）有限公司紧抓国家‘一带一路’建设契机，深入探索 AI 大数据在语言科技领域场景化应用的重要成果。同时，也将积极驱动中译语通陕西公司持续推进技术创新，充分利用大数据挖掘与分析、自然语言处理、语义计算、知识图谱、可视化分析等行业先进技术，深度布局金融、科技、智能制造、智慧旅游、‘一带一路’建设、公共舆情等诸多垂直领域，为全球企业级用户提供全方位的领先大数据与人工智能场景化应用解决方案，用科技助力‘一带一路’共建国家和地区之间经贸合作、人文交流的创新发展。”

【案例点评】

“一带一路”语言服务标准体系及大数据平台建立以来，面向政府和企事业单位、外向型企业、个人提供了有效的服务，显示出巨大的社会和经济效益。该平台为陕西乃至全国培养了一批掌握“关键语言”的外语人才，建立起“语言互联网”，大大提高了民众个人语言能力。此外，该平台还为陕西自贸试验区企业“走出去”提供基础的多语言服务保障，消除企业海外投资的各项信息壁垒。