

第三批“数据要素×”典型案例 ——商贸流通领域

多语言数据融合利用 助力行业高质量发展

语言数据存在标准不一、高度分散等问题，大量资源处于闲置状态，价值释放不足。传神语联网网络科技有限公司通过语言信息数据智能组织调度等技术研究，开发了语言区块链、数字孪生译员引擎、“任度”大模型等核心产品，实现了语言服务数据资源规模化和智能化的组织调度，推动行业生态信任机制构建，助力数据融合创新。

一是建设语料数据库，提升数据复用效率

构建语言服务行业数据采集、挖掘和应用体系，融合基础语言数据、专业语言数据、行业知识数据等相关数据，打造多语言语料数据库，并开放给行业企业和上下游合作伙伴。优化数据组织和调度，积累了海量语料及场景要求、过程数据、修改结果等信息，开发平台智能组织调度技术，提升语言服务综合能力。

二是基于区块链技术，推动语料数据流转

利用去中心化数据对象存储技术，梳理译员、语言服务公司与其语料数据资源的对应关系，保障各方权益。利用智能合约、行业知识资产交换等方式，推动语料数据资源安全有序流转。

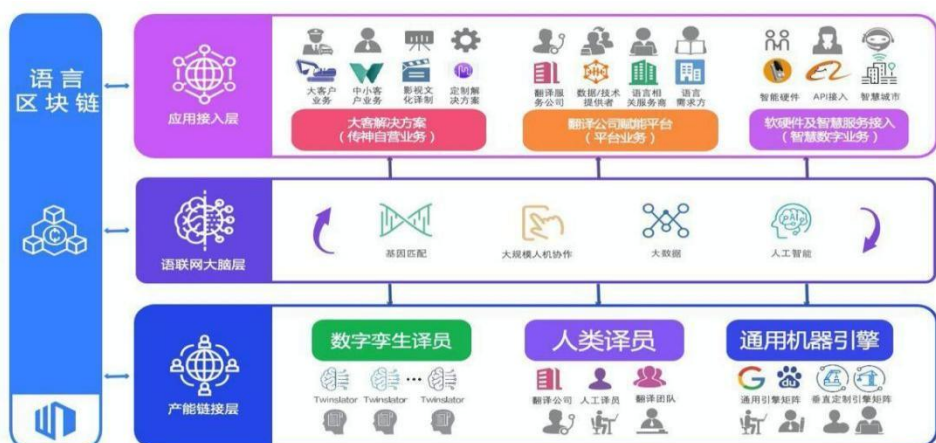


图 传神语联网推出的语言服务产品

三是开发大模型，推动数据融合创新

依托多语言语料数据，开发“任度”大模型，利用智能数据标注平台开展辅助标注业务，实现多模态数据的规模化生产与精准服务。融合行业知识，将“任度”大模型进行预训练，形成服务行业的垂直大模型，提升面向行业和场景的语言服务能力。

四是采用多重加密技术，保障数据安全

利用多重加密技术，对平台数据实施静态存储加密与动态传输加密。建立访问权限体系，精确分配数据读取、下载、修改等权限。将备份数据存储在异地数据中心，抵御硬件故障、自然灾害、人为攻击等风险，筑牢数据安全防线。

目前，语联网平台覆盖 80 多个常用语种、50 多个行业领域，收录近十万名译员行为数据、翻译过程数据，形成了超过 210TB 的多语言数据集资源。与 2000 余家企业达成合作，完成 2 万名语言服务人才信息上链，形成数万个数字孪生译员；与 150 余所高校形成合作，累计完成 5 万名学生的教育实践和能力测评。

数据共享融通助力跨境电商（“带电”货物） 运输效率提升

随着跨境电商货物在我国外贸中的占比持续攀升，含锂电池跨境电商商品的运输需求日益旺盛。然而，跨境电商锂电商品面临电池报告核验难题，导致跨境运输成本高昂、流通不畅。全球跨境电商“三中心”、广州云湾协同南航物流、广东机场物流等单位，联合开发了“全球跨境电商‘带电’货物数据服务平台”，为跨境电商“带电”货物在国内机场的便捷运输提供解决方案。

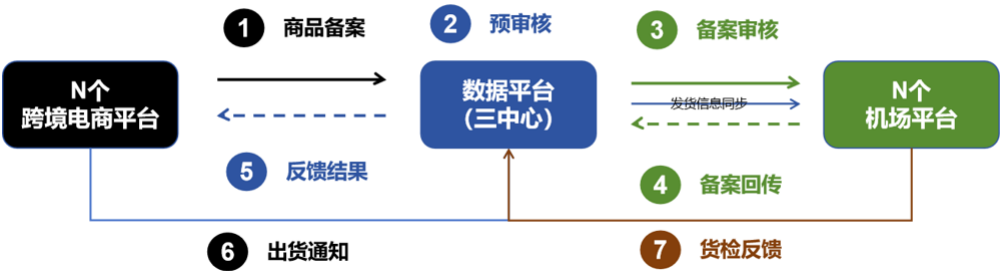


图 跨境电商“带电”货物平台数据对接流转设计图

一是汇聚“带电”货物相关数据

服务平台汇聚锂电备案、提单出货、货物清单等电商平台数据，收运检查、安检过机、装载飞行等机场平台数据，以及备案核对、真伪核验、贴标清单等自身平台数据，实现多源数据融合。

二是建立模型对数据进行审核

基于服务平台数据，搭建基础标准预审模型、商品信息风控模型、产品电池校验模型等安全风控模型。利用图像理解大模型等，对跨境电商“带电”货物进行智能识别，提取关键信息，并

结合商品备案信息，对“带电”货物相关数据进行预审核。

三是推进“带电”货物数据流通共享

依托服务平台，建立各电商平台与各机场之间的数据共享通道，确保信息的顺畅流通与认证，推出“一次预审、多端共享、快速出货”的高效便捷通关模式。

自平台建立运营以来，成效显著。提升运输时效性，佛山等地仓库到机场的运输时间缩短 5 小时，卡车运输时效提升 55%。成本控制优化，内地地面操作成本较香港降低 61%。提升用户体验感，本地货物出运效率提高，减少包裹拆分需求。

零售行业数据融合应用 赋能行业降本增效

零售业作为国民经济的重要组成部分，数据驱动的精准决策与流程优化，能显著提升其供应链敏捷性、服务个性化及运营智能化水平，增强市场竞争力，推动商贸流通行业降本增效。然而，实体零售行业数据开发利用仍面临数字化基础薄弱、企业认知不足、数据流通存在壁垒等多重挑战。多点生活（中国）网络科技有限公司构建“技术+咨询”双轮驱动模式，融合全链路数据管理与 AIoT 智能技术，为零售企业提供数智化解决方案。

一是采集汇聚零售全场景数据

通过 API 接口、IoT 设备，依法依规采集 SKU、价格、库存等商品信息数据，会员画像、购物路径等消费者行为数据，物流时效、仓储周转等供应链动态数据，以及能耗、设备状态等门店运营数据，日均处理数据量超 10 亿条。

二是开展数据分析建模

运用大数据分析和人工智能技术，构建智能补货模型，帮助商超、门店等零售企业精准预测商品需求，降低缺货率。通过分析消费者行为数据，帮助门店优化商品陈列策略，提升门店转化率。基于脱敏数据，定期生成月报及节日专报，为政府部门监测消费市场提供科学化、精准化的数据支持。

三是强化数据安全治理

为确保跨主体、跨层级流通过程中的数据安全，建立了数据

全生命周期防护机制。采用脱敏与匿名化技术，确保数据采集、存储、传输及使用安全。制定数据交互标准，为政企数据互通提供支撑。

目前，已服务 591 个客户，业务覆盖 10 个国家和地区，助力零售企业实现降本增效与营收增长。以物美超市为例，全面切换至 DMALL OS 系统后，订单增长率达 70%，会员人数增长超 50%，库存周转天数从 25 天缩短至 16 天，人效提升 35%。同时，通过智能结算、AI 导购等技术应用，大幅优化消费者体验，推动供应链与物流规划的智能化升级。

学术出版数据应用助力出版行业生产质量与效率提升

传统出版行业存在信息化程度低、内容形态单一等挑战，带来了生产效率低、成本高、周期长等问题。新译信息科技（深圳）有限公司通过系统整合学术出版数据资源，打造象胥编辑平台，提供智能化编辑服务，推动出版行业流程优化与改进，提升生产质量与效率。

一是整合汇聚学术数据资源

平台通过版权授权、数据加工采购等市场化方式，整合了 8500 亿字的优质文本资源，涵盖期刊论文、高校硕博论文、学术专著等权威学术资源，以及超 20TB 的双语平行语料库。

二是创新数据流通机制

构建“出版机构-高校-科研院所”三方协同的数据流通机制，联合出版社机构，制定《学术出版智能审校标准》，规范数据清洗、标注、传输等全流程，保障各方的数据权益，促进数据跨主体、跨层级流通。

三是打造强大智能编译服务

平台支持超百种语言（含外语及少数民族语言）的智能预翻译和外语审校，近两年 AI 翻译量突破 10 万亿字符。依托平台强大审校能力，实时捕获并标注字词语法错误、格式不规范等典型错例，形成“学术出版错例词库”，为编辑节省 50% 的审校时间。一键式智能处理实现投稿后自动润色、术语标准化与参考文献格

式化，将图书审稿周期从数月压缩至数周。依托 AI 翻译引擎与人工协作机制，使中英版本实现同步发布，将海外市场新书上市速度提升 60%。构建“AI 初筛-编辑复核-专家终审”三级质检机制，将科技专著排版错误率从 0.5‰降至 0.08‰。

四是强化数据安全保护

应用全链路加密、AI 安全监测、隐私保护计算等数据安全技术，保障数据跨主体、跨层级流通安全。自建数据中心，部署分布式存储集群与灾备系统。联合网络安全企业建立“白帽子”攻防实验室，定期开展渗透测试与红蓝对抗演练。

象胥编辑平台通过数据和技术赋能出版行业，优化出版流程，外语书籍采编效率提升 70%，年翻译文献突破百万篇，覆盖 30 多种语言。当前，平台已服务上百家出版机构，1300 余所高校通过平台获取多语种文献资源，每周完成 1 万篇外文文献翻译，助力科研人员节省 60%的文献获取时间。