

# 语言智能产业集聚与海南封关政策下数字经济协同发展研究

付鹏

三亚学院，海南三亚，572000；

**摘要：**本文探讨全球语言智能产业的发展，在当下海南自贸港数字经济发展的影响下，特别是海南封关运作后，语言智能产业如何结合海南自贸港政策发展语言智能产业，并协同数字经济协同发展进行研究。结合海南封关政策赋能机制，剖析研究政策创新点和智能产业的承接基础，研究语言智能技术和当下智能产业集聚区的发展现状，在构建“政策-技术-产业”三维分析框架下，为海南自贸港数字服务贸易提供路径参考。该论文将从语言核心技术图谱与语言智能产业集聚两大维度，详细剖析当前语言智能产业技术的发展现状。通过研究语言智能聚集下的各产业，结合海南封关后的各类政策，探讨语言智能+产业（包括语言智能+跨境电商，语言智能+国际教育，语言智能+智慧旅游等）的数字经济发展路径，建立和形成比较系统的智能产业集聚区。

**关键词：**语言智能；海南自贸港；封关；产业集聚

**DOI：**10.64216/3080-1486.25.07.003

## 引言

在当前全球化与智能化交织的大潮中，人工智能成为全球科技竞争的新焦点和经济发展的重要引擎。以人工智能大模型为代表，在数字技术方面的突破，推动了全球语言智能产业规模的迅速发展。语言智能产业规模的增长，对语言数字化的建设升级的提出强烈需求，推动了语言智能+产业的快速应用和市场渗透。特别是在海南自贸港“封关运作”（2025年底），基于《海南自贸港法》对跨境流动政策，以及国家发改委《数字经济国际合作指导意见》对语言智能的定位，旨在为全球数字贸易中的跨境数据流动治理提供一个深入而全面的研究视角，并为数字经济共同应对数字流动下的智能产业提供更多的机会。本文的研究意义在于构建“政策-技术-产业”三维分析框架，为海南封关后的自贸港数字服务贸易提供路径参考和时间价值。

## 1 海南封关政策的关键赋能机制

### 1.1 海南封关政策创新点

《海南自由贸易港数据跨境流动安全管理试点白名单管理办法》中显示，2024年Q1已有12家语言服务企业进入白名单，数据跨境“白名单”制度中，首批试点领域涵盖语言数据处理、金融征信等5个重要领域（海南省商务厅，2023）。数据显示，白名单内的企业数据出境审批时间大大缩短，由常规流程20日缩短至7个工作日内。数据跨境制度，大大减少跨境企业的数字处理速度和效率。二者，封关后增值电信业务对外开放政策有所放宽，《关于支持海南自由贸易港建设增值电信业务开放试验区的通知》中，增值电信业务外资股的占比放宽至100%。由此建立的国际互联网数据专用通道，

跨境通信性能指标下，延时大大降低（海南省通信管理局测试报告，2024）。这使的跨境企业之间的通讯成本大大节省，同时让语言服务跨国数据传输成本大大降低，对语言智能产业的生态发展提供数据支持，据海南生态软件园测算，语言服务企业跨国数据传输成本降低18-22%。<sup>[1]</sup>

### 1.2 语言智能产业承接基础

2024年数字产业经济年度报告显示，海南生态软件园语言产业集聚，为语言智能产业承接建立良好的基础，海南生态软件园目前已入驻语言服务企业43家。截至2024年6月，园区入驻语言服务企业43家（含科大讯飞、传神语联等头部企业），占海南省行业企业总数的68%（海南省工信厅数据）。对比2023年园区语言服务产业，同比增长比例高达42%，其中跨境语言数据处理占比35%。<sup>[2]</sup>与此同时，三亚崖州湾科技城多语言服务能力显著提升，三亚崖州湾科技城建设多语言呼叫中心，基建投入资金1.2亿元用于建设多语言呼叫中心，支持英语、法语、俄语等8种语言服务（三亚市发改委批复文件，2023）。为海南封关后的语言智能产业的发展提供技术支持参考。

## 2 语言智能技术发展

语言智能产业作为人工智能(AI)领域的重要分支，特别是在海南自贸港政策推动下呈现出蓬勃发展态势。当下热门的chatGPT和Deeseek等相关语言智能软件的加持下，语言智能化发展呈现前所未有的变革状态。语言智能技术发展，包括语音识别、机器翻译、情感计算等关键性技术的突破与应用，<sup>[3]</sup>本文将详细阐述如下：

### 2.1 语音识别技术

语音识别技术作为语言智能技术的重要应用，在人机交互的技术层面下，已实现从通用场景到各个细分领域下的纵深发展。当前主流语音识别系统在安静环境下的普通话识别准确率普遍超过 95%，而方言识别成为技术攻关的重点方向。例如：科大讯飞已成为中国语言智能技术行业的领军企业，其方言识别引擎技术已深度适配部分方言，例如壮语、桂柳话等。科大讯飞的语音识别技术在复杂环境下的识别准确率达到 92%。同时，软件支持普通话、英文与多种方言的混合输入，在多方言混合 ASR 模型的发展，使口音识别准确率达 92%。此技术在市场应用发展方面，已渗透至服务产业、医疗产业、教育产业等多个垂直领域。

### 2.2 机器翻译技术

机器翻译技术是语言智能技术的核心要素，基于早期的翻译规则和翻译统计方法，更新迭代到以神经网络为主导的智能翻译阶段，以 Google 的 Transformer 架构为例，Transformer 系统通过“关注机制”(Attention Mechanism)，配对比较段落中的每个单词，通过关联上下文，提高翻译内容准确性。特别是在“英文→德文”和“英文→法文”的翻译任务中，显示出强大的领域适应性。现代机器翻译技术，系统面临的核心挑战是多义词理解和长距离依赖问题。以“bank”单词的机器翻译为例，传统逐字翻译算法难以确定其指代“银行”还是“河岸”，而 Transformer 的关注机制能够通过上下文并分析单词间的相关性进行评分，Transformer 较早期的翻译能够更加准确的捕捉上下文线索。<sup>[4]</sup>同时在机器翻译技术的发展在语料库的建设中，AI 基地的语料库共享建设和数据交易为专业领域翻译提供强有力的技术支持。

### 2.3 情感计算技术

情感计算技术作为语言智能的高级应用，通过用户

集聚区	核心优势	企业数量	典型企业	政策支持
北京中关村	高校科研资源密集	约 2000 家	深度求索、智谱 AI	国家语料库建设
深圳南山	硬件配套完善	约 1500 家	腾讯、平安科技	"语料券"政策
粤港澳大湾区	多语言环境	约 3755 家	腾讯、科大讯飞	跨境数据流通试点

北京中关村作为技术研发型集聚区，拥有北京大学、清华大学等顶尖高校及中国科学院等各类科研机构，高校科研资源密集分布，典型的企业深度求索为代表，拥有自然语言处理基础算法、大语言模型等核心技术，地理优势在智能科技研发上具有明显优势。教育部已布局新型国家语料库建设，其重点支持多语言智能化科技创新中心和各类语言实验室建设，为中关村企业提供了强大的科研支撑。

相比之下，深圳南山则依托比较完善的硬件产业链，

的情感体验数据，建立客户服务、心理健康等领域的情感交互体验。目前大多数人工智能技术在大部分语言客服系统应用中，能够实现比较高的的情感识别准确率。在技术创新方面，现代情感计算引擎已从单一文本分析发展为多模态融合系统。例如苹果 Vision Pro 的多模态传感器系统可通过摄像头，通过捕捉用户唇部运动轨迹、检测面部肌肉微小振动，结合眼动追踪，实现高精度的情感识别，在实验室环境下情感识别准确率可达 92%。这种“视觉+触觉+环境感知”的三重融合技术正在重构人机交互范式，成为当下情感技术支持语言智能技术的重要基础。情感计算显著提升了服务行业的智能化水平。人工智能语音客户服务系统通过动态情感计算，可以实现从基础语音交互模式到深度业务理解的变革，例如在心理健康的非接触式监测领域，多模态情感分析技术可以通过语音特征、微表情、眼动追踪等指标，实现抑郁、焦虑等心理状态的早期筛查。

文中所提及语言智能产业的三重技术，从语音识别技术的发展，到机器翻译技术人工智能化的更新迭代，连同情感计算技术的建立，很好的推进语料库、数据标注、数据评价等规范标准的修订，为人工智能高级语言应用奠定良好的基础。

## 3 语言智能产业集聚发展

语言智能产业受地理空间和政策影响下呈现比较明显的集聚效应，本文以北京中关村的技术研发和深圳南山硬件发展，以及粤港澳大湾区的智能产业发展为例，研究分析其产业集聚性，其共同性都是凭借差异化的地理优势引领智能产业的发展。接下来将阐述企业产业集聚区的差异化发展路径，同时结合典型企业案例与最新数据，全面呈现这一领域的创新生态环境与有利的市场格局。下表为语言智能产业集聚区截止 2025 年 6 月统计的中国主要语言智能产业集聚区的比较：<sup>[5]</sup>

已发展成为智能产业硬件配套型集聚区。深圳市工业和信息化局印发的《深圳市加快打造人工智能先锋城市行动计划(2025-2026 年)》提出，充分利用深圳产业发展优势及“语料券”政策，持续构建行业语料集，更新语料库建设。同时深圳数据交易所目前正在打造国家级数据交易所，发放“语料券”的同时，促进人工智能语料的共享和市场交易，为语言智能企业提供了丰富的数据资源和完善的产业配套设施。

另外，粤港澳大湾区作为语言智能产业的重要集聚

地,已形成“广深双核+次级节点”的空间结构。数据显示,2024年大湾区有3775家语言数据标注相关企业,实现产值44.02亿元,拥有从业人员1.28万人。深圳和广州聚集了88.24%的企业,贡献95.47%的产值。港澳特区产业建设虽直接参与有限,但在跨语言数据标注中发挥独特作用,特别是中国澳门中西创新学院与北京语言大学联合发布的《粤港澳语言数据标注产业发展白皮书(2025)》为区域产业发展提供了重要指南。<sup>[6]</sup>

#### 4 海南封关政策下语言智能协同数字经济协同发展路径

语言智能产业的健康发展离不开政策引导和标准规范。在国家层面,教育部、国家语委、中央网信办联合印发的《关于加强数字中文建设推进语言文字信息化发展的意见》明确提出“两步走”发展目标:到2027年形成语言文字信息化推进机制;到2035年使我国语言文字信息化整体水平位居世界前列。除此之外,海南省封关下对语言智能系统的核心政策提供良好基础,为下一个智能产业集聚区提供帮助,相关政策如下:

1. 语言服务产业支持政策:根据《海南省国际智能语言服务产教融合共同体建设》(2025年)提出,聚焦智能翻译技术、跨文化传播、涉外语言服务三大领域,推动“政产学研用”资源整合,建立智能翻译技术研发和人才培养平台。政策支持高校与企业协同开发多语种语料库、AI翻译引擎,并服务于海南自贸港的国际会展、旅游等场景。

2. 数据要素开放与跨境流动政策:《海南自由贸易港数据出境安全管理负面清单》(2024年试行)中的语言智能相关条款下,允许旅游、医疗等领域语料数据在安全评估后跨境传输,支持多语种语料库建设;金融、航运领域纳入“跨境数据流动白名单”,开放多语言客服对话数据用于AI训练。<sup>[7]</sup>

3. 智能语言应用场景拓展政策:《海南自由贸易港口岸管理服务条例(试行)》(2025年12月18日施行),要求建设“智慧口岸”,支持多语种智能交互系统在通关服务中的应用,例如多语言报关指引、信用提示等。《海南自由贸易港“二线口岸”通关信用管理若干规定》下,明确将语言技术服务纳入“高级认证企业”加分项,例如提供多语种合同审核、跨境仲裁语言支持的企业可获通关便利。<sup>[8]</sup>

封关政策为跨境服务、国际教育和智慧旅游提供强大的政策支持,海南将建立和形成新的智能产业集聚区,为语言智能+产业(包括语言智能+跨境电商,语言智能+国际教育,语言智能+智慧旅游等)的数字经济发展提供切实可行的建设路径。海南封关对语言智能产业发展

提供各类支持,海南的语言智能政策呈现“技术研发—数据开放—场景应用”的三层支持体系,即智能科技研发的高品质研发,数据资源的开放政策,各类场景多应用的支持系统;教育部布局的新型国家语料库建设,突破传统语料库建设,进行语言资源整合;跨境数据白名单为跨境电商提供数字支持,多语种智能交互承接智慧场景的拓展,为企业在市场中的通关提供便利。

#### 5 结论

综上所述,海南自贸港数字经济发展在海南封关运作后各类政策支持下,海南岛建立和形成新的智能产业集聚区,语言智能发展资源丰富,跨境语言智能产业基于“政策-技术-产业”三维框架,为海南封关后的自贸港数字服务贸易提供路径参考和时间价值。使得语言智能产业技术的发展取得关键性技术的突破与应用,在各企业产业集聚区的差异化发展路径下,呈现智能科技领域的创新生态与秩序良好下的市场格局。根据海南封关政策语言智能呈现“技术研发—数据开放—场景应用”的三层支持体系,使得语言智能+产业(包括语言智能+跨境电商,语言智能+国际教育,语言智能+智慧旅游等)协同数字经济协同发展,并为构建丰富的跨境语言智能文化环境做出贡献。

#### 参考文献

- [1]康欣.数字贸易中跨境数据流动国际规制[D].北京外国语大学,2024. DOI: 10.
- [2]工业和信息化部.(2024).关于支持海南自由贸易港建设增值电信业务开放试验区的通知(工信部信管〔2024〕12号).
- [3]钟俊浩,赵春昊,饶雪.人工智能标准体系:夯实产业发展基石的“上海实践”[J].质量与标准化,2025,(03):10-12.
- [4]Google Research.(2024).“Transformer-XL in Low-Resource Languages”.
- [5]张薇.(2024).“语言智能产业集聚效应研究”.《科技管理研究》,44(3).
- [6]海南统计局.(2024).《数字经济产业年度报告》.
- [7]海南省商务厅.(2023).海南自由贸易港数据跨境流动安全管理试点白名单管理办法[EB/OL]
- [8]海南省人大.(2021).《海南自由贸易港法》.

作者信息:付鹏,女,1984年1月出生,汉族,籍贯河北,悉尼科技大学商科硕士,现任三亚学院教师,研究方向为国别与区域研究。