

广西面向东盟的大模型技术应用路径研究

随着人工智能技术发展，大模型已成为数字经济发展的核心赛道。东盟小语种资源丰富但面临语料匮乏、算力不足等瓶颈。广西作为中国面向东盟唯一陆海双通道，依托东博会等平台，叠加跨境贸易与语料禀赋，具备深度参与区域大模型合作的独特优势。建议广西以技术协同研发与轻量化部署为核心方向，以电力、算力保障为基础支撑，以“小语种+AI”人才培育为动能，以政策互认与信任赋能为保障，四大举措并举探索大模型跨境合作路径，推动中国及广西大模型技术与产品在东盟规模化落地应用。

一、东盟大模型发展态势与核心特征

（一）小语种大模型成区域数字竞争核心赛道，发展环境差异明显

东盟有1200种语言资源¹，凭借独特的语言多样性禀赋，已成为小语种大模型研发核心热区。外部布局层面，呈现单向主导和开放合作两种模式。美国科技巨头以高资本投入、高技术壁垒、高规则主导，构建以自身为核心的东盟大模型布局体系，ChatGPT Go 以多语言通用性与技术成熟度占据先发优势，已推广至东盟十国；中国则以技术转让、联合研发、开源共享为核心路径，凭借开放兼容、合作共赢的技术

¹ 数据来源：大力财经《谷歌 AI 攻克亚洲语言难题，2300 种语言数字化计划正在推进》

优势，在研发成本与性能上实现突破，阿里 SeaLLM 东盟模型通过分词优化实现降本 50%²，科大讯飞“星火东盟多语言大模型”攻克 7 种东盟语言实时翻译技术³。区域内部发展层面，合作赋能需求与安全保障要求并重。新加坡基于阿里 Qwen 基础模型研发 SEA—Lion，完成混合方言适配，可精准理解本地语境、实现低门槛部署；马来西亚、泰国以技术引进与本地化适配为主，采取轻量化适配与安全合规并重模型发展路线，重视语言文化保护与垂直场景应用；越南、印尼兼顾技术自主与合规应用，聚焦本土语言适配与跨境贸易、智慧农业等垂直领域场景落地；老挝、柬埔寨、缅甸、文莱等依托跨境合作补短板，聚焦基础语言与民生场景，推进本土化大模型的研发迭代与应用探索。

（二）发展资源匮乏，政策规则多样化

一是语料资源不足，研发成本高，场景适配不足。东盟小语种语料匮乏且错漏率较高，模型准确率普遍不足 70%⁴，开发成本约为大语种模型数倍⁵；其中美国科技巨头模型训练及推理所需算力成本高，对东盟本土方言适配欠缺且实体产业场景赋能不足；新加坡本地自研模型参数规模受限，对产业场景适配能力不足，技术框架依赖外部，云与应用生态支撑薄弱。二是算力需求与电力供应短缺矛盾突出。东盟农村互联网覆盖率仅 42%⁶，数据中心渗透率较中美等较发达市场

² 数据来源：《SepLLM: Accelerate Large Language Models by Compressing One Segment into One Separator》

³ 数据来源：新浪财经《今天 AI 唱主角！科大讯飞发布星火东盟多语言大模型底座及系列 AI 产品》

⁴ 数据来源：昆明理工大学学报（自然科学版）《东南亚低资源语言神经机器翻译研究进展》

⁵ 数据来源：牛津大学《Language Model Tokenizers Introduce Unfairness Between Languages》

⁶ 数据来源：广西壮族自治区商务厅《中国—东盟电子商务发展报告》。

落后 55%—70%⁷，且区域电力基础设施薄弱、可再生能源支撑不足，难以匹配模型算力设施规模化部署需求。三是人才供给存在结构性缺口，AI 人才缺口达数十万⁸，“小语种+AI”复合型人才尤为稀缺，严重制约模型技术发展。四是东盟政策多样规则适配难。各国数据跨境流动与模型跨国部署的合规标准不一，适配成本高；区域内语言文化、伦理价值差异显著，模型应用易引发文化误读与价值观冲突；数据安全约束下跨境信任壁垒持续存在。

二、广西发展面向东盟大模型的现实基础

区位与平台优势显著。广西是中国面向东盟唯一陆海双通道，中国—东盟自贸区 3.0 版、RCEP 生效及平陆运河建设强化枢纽地位；广西跨境陆缆总容量超 9000GB⁹，至越南时延仅 3.5 毫秒¹⁰。东博会 AI 专馆促成超 140 项成果、成交额 1.4 亿元¹¹；中国—东盟国家人工智能应用合作中心吸引 50 余家企业入驻，部署算力 5000PFlops¹²，AI 赛事吸引 584 支东盟队伍¹³，加速形成 AI 创新集聚与交流载体，推动区域模型协同创新。贸易与产业基础扎实。东盟连续 25 年为广西第一大贸易伙伴，沉淀海量边民交易等经贸、文化语料，形成区别于 SEA—Lion 的东盟本地数据优势。广西企业深耕

⁷ 数据来源：马来亚投资银行 2025 年发布的《ASEANDataCentre:Keepcalmandcarryon》

⁸ 数据来源：2025 年《AI 在东南亚驱动商业增长》

⁹ 数据来源：金台资讯《广西建设面向东盟的“数据流通便利地”》

¹⁰ 数据来源：新华网南宁 9 月 20 日电《中国—东盟国家人工智能应用合作中心：造中国—东盟人工智能合作新枢纽》

¹¹ 数据来源：南宁市融媒体中心《进馆人数创历史新高！第 22 届东博会圆满落幕，实现人气与实效双高》

¹² 数据来源：新华网南宁 9 月 20 日电《中国—东盟国家人工智能应用合作中心：打造中国—东盟人工智能合作新枢纽》。

¹³ 数据来源：广西日报《潮涌 AI 智联东盟—广西“A 超”链出协同硕果》。

东盟当地，中国产业 AI 技术可通过广西实现东盟本地化复制，形成“中国产业 AI+广西集成+东盟应用”链条。市场需求与中国生态共振。东盟在智慧农业、跨境物流等领域的数字化转型需求迫切，与广西的垂类模型研发主攻方向高度契合；中国正构建大模型开源协同生态闭环，通过技术共享、多元协作与自主迭代能力，可助力东盟规避单一技术依赖风险、保障数字经济自主可控运转，显著提升对东盟的吸引力。

三、面向东盟的大模型研发与应用路径优化建议

（一）聚焦协同研发与轻量化部署方向，攻克小语种模型性能及本土化适配痛点

联合科研机构与北上广龙头企业共建小语种模型实验室，推进模型开源化、轻量化、本地化、安全化部署。搭建模型公共服务平台与开源社区，构建组件化、可复用的成熟模型资产库，构建“通用+垂直”应用矩阵。以开源模型为基础适配小语种，支持断网运行与低带宽本地化部署，降低落地应用门槛，扩大算力薄弱区域应用覆盖范围，并适配其数据安全性与文化伦理要求，实现模型低成本供给、低门槛使用、多场景适配，助力产业升级。

（二）强化电力、算力保障，筑牢东盟大模型产业硬件根基

一是构建绿色算力与跨境电力协同体系。推广液冷等绿色技术，提速核电、新能源与算力一体化中心建设及智算中心扩容，依托园区专项补贴优化电价、降低用算成本。聚焦东盟电力供给缺口与算力增长需求，探索推出“基础电力+

弹性算力”一体化打包服务方案，配套提供运维技能培训与全流程技术支持，助力东盟加速电力与算力基础设施升级提质。二是构建跨境算力与本地化适配协同服务体系。探索搭建“国内超算—广西枢纽—东盟边缘”三级算力调度平台，破解算力供给难题。构建模型与芯片一体化算力底座，联动东盟本土企业完善跨境运维服务、本地化技术适配、供应链协同保障等全链条配套，破解单一技术依赖。

（三）激活“小语种+AI”人才育引留动能，缓解人才供给制约

一是校企协同育才赋能产业。依托“中国—东盟博览会教育展”“跨境教育合作联盟”等平台，每年举办专业共建对接会，推动双方高校增设“AI+小语种+东盟产业场景”交叉专业，构建适配课程体系，联合企业共建多语种语料库与实训基地，实现人才培养与东盟产业需求精准匹配。二是精准引才补齐发展短板。出台住房补贴、科研经费配套等政策，通过联合攻关、短期访问等柔性模式，为东盟引进“小语种+AI”高端人才及团队。三是强化人才留存。推动技能证书联合认证，联合实验室开展实践项目，孵化 AI 产业，打造“培养—就业—晋升”岗位闭环，筑牢人才留存保障。

（四）推动政策互认与信任赋能，破除合作壁垒

一是搭建制度互认框架。衔接《东盟数字经济框架协议》（DEFA）、《东盟 AI 治理与伦理指南》《东盟数字数据治理框架》，依托数据跨境流动试验区及中国—东盟国家人工

智能应用合作中心等平台，协同制定数据标注规范、算法伦理准则及安全评估标准等互认型规则，在低敏感领域试点数据跨境流动“白名单”机制，为区域 AI 大模型产业合作筑牢政策互认制度根基。二是强化互认实践与信任绑定。联合共建行业场景库与解决方案库，依托东博会打造“需求对接—方案匹配—效果验证—互认推广”闭环，以合作实践积累互认经验；系统梳理互认成果形成合作报告，在东盟城市设立展示体验中心，公开技术成效与安全措施，夯实政策互认的社会与市场基础。

（执笔人：付洁、满彦星、梁颖、文进、罗传）