E-Commerce Letters 电子商务评论, 2026, 15(1), 978-990
Published Online January 2026 in Hans. https://www.hanspub.org/journal/ecl
https://doi.org/10.12677/ecl.2026.151118

Hans汉斯

电子商务与大模型融合发展的核心瓶颈及解决路径研究

王枭宇

贵州大学省部共建公共大数据国家重点实验室，贵州 贵阳

收稿日期：2025年12月22日；录用日期：2025年12月31日；发布日期：2026年1月30日

摘 要

随着ChatGPT、文心一言等大模型技术的迭代成熟，其在电子商务领域的应用已从早期试水进入规模化价值兑现阶段，在智能推荐、供应链优化、用户服务等环节展现出显著赋能效应。然而，技术融合过程中暴露出的适配性不足、数据治理失序、商业落地受阻、伦理合规缺失等问题，严重制约了融合价值的充分释放。本文基于文献梳理与行业案例分析，系统识别电子商务与大模型融合过程中急需解决的四大核心问题：技术层面的场景适配与泛化能力不足、数据层面的质量管控与合规风险凸显、商业层面的中小企业落地门槛与价值兑现困境、伦理层面的算法偏见与内容可信性危机。在此基础上，结合最新技术进展与政策框架，从技术优化、数据治理、商业赋能、伦理监管四个维度提出针对性解决路径，构建“问题识别–成因分析–路径构建”的完整研究链条，为推动电子商务与大模型深度融合提供理论支撑与实践指引。

关键词

电子商务，大模型，数据治理，算法伦理，全生命周期治理

Core Bottlenecks and Solution Paths for the Integrated Development of E-Commerce and Large Language Models

Xiaoyu Wang

State Key Laboratory of Public Big Data, Guizhou University, Guiyang Guizhou

Received: December 22, 2025; accepted: December 31, 2025; published: January 30, 2026

文章引用: 王枭宇. 电子商务与大模型融合发展的核心瓶颈及解决路径研究[J]. 电子商务评论, 2026, 15(1): 978-990.
DOI: 10.12677/ecl.2026.151118

Abstract

With the iterative maturation of large language models (LLMs) such as ChatGPT and ERNIE Bot (Enhanced Representation through Knowledge Integration Bot), their applications in the e-commerce field have evolved from the early pilot phase to a stage of large-scale value realization, demonstrating significant enabling effects in links including intelligent recommendation, supply chain optimization, and user services. However, problems exposed in the process of technological integration—such as insufficient adaptability, disordered data governance, hindered commercial implementation, and lack of ethical and regulatory compliance—have seriously restricted the full release of integrated value. Based on a literature review and industry case analysis, this paper systematically identifies four core issues urgently requiring resolution in the integration of e-commerce and LLMs: insufficient scenario adaptability and generalization capability at the technical level, prominent quality control and compliance risks at the data level, high implementation thresholds for small and medium-sized enterprises (SMEs) and value realization dilemmas at the commercial level, and algorithmic bias and content credibility crises at the ethical level. Building on this, integrating the latest technological advancements and policy frameworks, the paper proposes targeted solutions from four dimensions—technological optimization, data governance, business empowerment, and ethical supervision—and establishes a complete research chain of "problem identification—cause analysis—path construction," providing theoretical support and practical guidance for promoting the in-depth integration of e-commerce and LLMs.

Keywords

E-Commerce, Large Language Models (LLMs), Data Governance, Algorithmic Ethics, Full Lifecycle Governance

1. 引言

1.1. 研究背景

自 2022 年 ChatGPT 发布引发全球 AI 热潮以来，大模型凭借其多模态处理、泛化推理与内容生成能力，成为重构电子商务生态的核心引擎[1]。从头部电商平台的智能搜索优化到中小卖家的自动化运营工具，大模型已渗透至电商“人货场”全链路：丽人丽妆通过虚拟人直播覆盖 40%直播时长，大幅降低运营成本；焦点科技 AI 麦可 4.0 实现外贸工作流自动规划，产品曝光量提升 21.5%；淘宝“千牛 AI 助手”针对国内中小卖家优化方言适配与节日营销场景，使中小卖家运营效率提升。国金证券《电商行业深度报告：AI + 电商服务进入提效阶段，关注后续业绩兑现》(2025)数据显示，2025 年电商行业 AI 应用已从“降本”向“提效”转型，成为驱动业绩增长的重要引擎[2]。

尽管融合发展前景广阔，但实践中仍面临诸多现实障碍。国际层面，亚马逊在印度市场的 AI 布局因本土语言适配不足、数据合规受限，其对话式搜索转化率较本土对手 Flipkart 低 37% [3]；国内层面，部分电商平台因大模型生成虚假促销信息遭监管处罚，抖音电商则通过完善算法备案流程、搭建内容审核机制，成为首批通过《生成式人工智能服务管理暂行办法》备案的平台之一[4]。这些案例表明，电子商务与大模型的融合仍处于“机遇与挑战并存”的关键阶段，梳理并解决其中的核心问题具有重要的现实

意义。

1.2. 研究现状

现有研究主要聚焦于大模型在电商领域的应用价值与技术实现，形成了三大研究方向：一是技术赋能层面，学者们关注大模型在个性化推荐、供应链预测、智能客服等场景的效率提升效果，证实其可使复购率提升 25%以上、库存周转率优化 60%，智能客服一次性解决率提升 30%~40% [5] [6]；二是模型优化层面，通过构建电商专属数据集(如 eCeLLM 的 ecInstruct 数据集)，提升大模型在电商场景的适配性，验证了“通用大模型 + 垂直精调”模式的优越性[7]；三是风险管控层面，部分研究提及数据隐私与算法伦理问题，提出采用联邦学习、数据脱敏等技术保障数据安全[1]，同时探讨了推荐系统流行度偏见等算法公平性问题的测度方法[8]。

现有研究虽揭示了融合发展的部分痛点，但仍存在明显不足：其一，对技术适配的深层矛盾(如多语言场景适配、冷启动问题)分析不够深入，缺乏对本土文化习俗、消费习惯与模型适配的关联性研究[9]；其二，缺乏对中小企业落地困境的系统性探讨，未关注“算力资源分配不均”“定制化解决方案缺失”等核心成因，尤其在跨境电商场景下中小企业的 AI 应用瓶颈研究较为匮乏[10]；其三，对合规备案、算法偏见等伦理合规问题的实证研究较为匮乏，现有研究多停留在理论层面，缺乏结合具体政策框架的落地性分析[4] [10]。基于此，本文聚焦“急需解决的核心问题”，构建技术–数据–商业–伦理多维度分析框架，弥补现有研究的不足。

1.3. 研究方法与研究内容

本文采用文献研究法、案例分析法与归纳演绎法相结合的研究方法：通过梳理国内外核心期刊(如《电子商务研究》《计算机应用研究》)、CSSCI 来源文献、权威行业报告(国金证券、艾瑞咨询)等相关资料，明确研究基础与理论边界；结合亚马逊印度布局、Flip kart 本土创新、淘宝“千牛 AI 助手”、抖音电商合规备案等典型案例，提炼实践中的核心问题；基于技术、数据、商业、伦理四大维度，归纳问题本质并提出解决路径。

研究内容聚焦四大核心问题(按影响优先级排序)：技术层面的场景适配与泛化能力不足(首要瓶颈)、数据层面的质量管控与合规风险凸显(基础约束)、商业层面的中小企业落地门槛与价值兑现困境(关键障碍)、伦理层面的算法偏见与内容可信性危机(生态保障缺失)，形成“问题识别–成因分析–路径构建”的完整研究链条。

2. 电子商务与大模型融合的核心问题识别

2.1. 技术层面：场景适配不足与泛化能力受限

2.1.1. 本土场景与语言适配性缺失

大模型的技术效果高度依赖本地化语料支撑，而电商场景的地域差异性加剧了适配难度。这种适配不足体现在三个维度：语言适配、文化适配与消费习惯适配。国际案例中，亚马逊在印度市场的 AI 工具因未能覆盖印地语、泰米尔语方言，且忽略“手工制作”“非遗认证”等本土消费标签，导致中小卖家库存周转效率提升仅 12%，远低于 Flipkart 的 45% [2]；国内案例中，早期部分跨境电商 AI 翻译工具因未适配小语种方言，导致商品描述翻译误差率达 28%，引发大量退换货纠纷[2]。文化适配层面，在宗教节日密集的印度市场，亚马逊的通用需求预测模型多次失误，而 Flipkart 的“节日需求预测系统”因整合本土宗教节日消费数据，精准预判需求峰值，成效显著；消费习惯适配层面，淘宝“千牛 AI 助手”初期因未考虑下沉市场卖家的低带宽使用场景，导致工具加载失败率达 15% [2]，后续通过轻量化优化才得以改

善。

2.1.2. 冷启动与泛化性矛盾突出

电商场景的“新用户、新产品”冷启动问题是大模型难以有效解决的核心痛点。传统电商模型因任务专一性，在未见过的产品或用户场景中表现不佳；尽管大模型具备一定泛化能力，但在电商领域仍存在明显局限 eCeLLM 的研究证实，现有大模型对 unseen 产品和 unseen 指令的处理效果仍有较大提升空间。其核心成因在于：冷启动场景缺乏足够的历史数据支撑，大模型的泛化能力难以突破“数据依赖”瓶颈。例如，新上线的小众设计师品牌因缺乏历史销售数据，大模型无法精准生成适配的营销文案，也难以实现有效的用户匹配；下沉市场新用户因行为数据匮乏，个性化推荐准确率较成熟用户数量低，导致用户留存率下降。

2.1.3. 提示工程与供应链场景适配失效

在电商供应链的需求预测、库存优化等核心场景中，大模型的提示工程存在显著局限性，根源在于大模型对电商供应链“动态复杂系统”的认知不足，仅依赖文本交互难以精准捕捉场景背后的业务逻辑。具体表现为两个问题：一是关键变量遗漏，实践中精心设计的提示常因忽略竞争对手新品推出、港口罢工、政策变动等外部因素导致预测失误；二是潜规则理解不足，部分模型输出的安全库存较实际需求高出 30%，原因在于未充分考虑“仓库容量限制”“物流时效波动”等供应链潜规则。例如，某头部电商的 AI 库存优化工具因未纳入“双十一前置仓备货规则”，导致大促期间部分前置仓库存积压、部分区域缺货，损失超千万元。为精准量化此类供应链场景下大模型提示工程的适配失效程度，可采用库存预测误差率公式进行衡量：

$$\text{库存预测误差率} = \frac{\text{预测库存需求量} - \text{实际库存需求量}}{\text{实际库存需求量}} \times 100\%$$

其中，预测库存需求量为大模型预测的库存需求量，实际库存需求量为实际市场需求对应的库存需求量，该公式可精准衡量供应链场景下大模型提示工程的适配失效程度，误差率越高说明适配性越差。

2.2. 数据层面：质量管控失序与合规风险凸显

2.2.1. 多源数据质量参差不齐

电商数据具有多模态、异构性特征，涵盖用户行为、交易记录、内容反馈等 12 类数据源(根据艾瑞咨询《2025 年中国电商数据生态报告》分类标准)，但其质量问题严重制约大模型效果。核心问题包括数据噪声、数据缺失与数据标注不规范。数据噪声方面，刷单数据、无效评论等污染训练语料，导致模型推荐准确率下降；数据缺失方面，中小卖家因数据采集能力不足，难以提供高质量训练数据，形成“数据匮乏–模型效果差–业务提升有限”的恶性循环；数据标注方面，亚马逊印度外包团队交付的 AI 项目 Bug 率是欧美团队的 4.6 倍，核心原因在于数据标注流程不规范、标注人员缺乏电商业务认知，导致标注错误率高。据行业调研显示，跨境电商 AI 项目因数据质量差异，本土团队与海外团队的 Bug 率差距可达 3~5 倍(调研范围：印度、东南亚跨境电商服务商，样本量 n = 23)。

2.2.2. 数据合规与跨境治理困境

数据跨境传输相关合规要求成为融合发展的重要约束，不同地区的法规差异进一步加剧了合规难度。国际层面，印度《数字个人数据保护法案》要求关键个人数据本土存储，迫使亚马逊耗费 15 亿美元进行技术改造，且因无法联动全球数据资源，其多语言模型方言识别准确率仅 68%；欧盟《通用数据保护条例》(GDPR)对数据跨境传输的严格限制，导致跨境电商企业的用户行为数据无法有效整合，模型训练效

果受影响。国内层面，我国《生成式人工智能服务管理暂行办法》明确要求训练数据合法合规，但实践中仍存在诸多难题：抖音电商在备案过程中，因部分数据脱敏不达标、来源证明缺失，备案材料曾两次被退回，产品上线周期延长 2 个月；部分中小跨境电商企业因缺乏合规团队，难以应对不同地区的法规差异，面临数据跨境传输的合规风险。

2.2.3. 大模型训练应用中的隐私泄露风险

大模型训练与应用过程中存在双重隐私泄露风险：敏感数据脱敏缺失风险与模型记忆引发的二次泄露风险。敏感数据脱敏缺失方面，电商数据包含大量个人敏感信息(如消费习惯、支付记录、收货地址)，若未经过严格脱敏处理，易引发隐私泄露——国内某电商平台因用户支付记录脱敏不彻底，导致 10 万条用户信息泄露，遭监管部门罚款 500 万元；亚马逊印度团队信息泄露率高，导致其失去多个欧洲品牌合作订单。模型记忆引发的二次泄露方面，大模型的“记忆特性”可能导致训练数据中的敏感信息被提取，例如，通过特定提示词可诱导大模型输出训练数据中的用户手机号、邮箱等信息。尽管联邦学习、差分隐私等技术已在部分企业应用，但因成本高、技术门槛高，尚未实现规模化普及，中小企业难以负担相关技术部署成本。

2.3. 商业层面：基于 TOE 框架的中小企业 AI 采纳障碍分析

TOE (Technology-Organization-Environment)框架是解释企业技术采纳行为的经典理论模型，其从技术特性(Technology)、组织内部特征(Organization)、外部环境因素(Environment)三个维度系统梳理采纳障碍。结合电商行业特性，本文基于 TOE 框架剖析中小企业大模型采纳困境，具体如下：

2.3.1. 技术维度(Technology)：适配性不足与资源约束

在技术维度(Technology)，技术自身的特性与资源依赖构成了中小企业采纳大模型技术的核心障碍，具体体现在场景适配性缺失、算力资源刚性约束及技术复杂度较高三个层面。

场景适配性缺失是首要问题。当前大模型服务商多聚焦头部企业需求，所提供的解决方案普遍存在“算法套用”现象，缺乏针对中小企业业务场景的轻量化、定制化优化，尤其在低带宽适配、方言交互等实用性功能上存在明显短板，导致技术与中小企业的实际运营需求脱节。例如，亚马逊印度推出的“智能定价系统”直接移植欧美市场成熟算法，未充分考量印度区域内显著的购买力差异(北方邦与南方邦购买力差距达 2.5 倍)，最终引发大量卖家抵触；国内某 SaaS 服务商的 AI 工具因未兼顾下沉市场低带宽使用场景，加载失败率高，难以满足中小卖家的核心业务需求。

算力资源的刚性约束进一步加剧了采纳难度。电商行业的算力需求呈现 “大促峰值波动”特征，峰值时段算力需求可达日常的 5 倍。头部平台可通过自有数据中心的规模化优势分摊算力成本，而中小企业普遍缺乏私有云部署能力，只能依赖公有云服务，但其弹性扩容费用在大促期间会增长 3~4 倍，远超中小企业的成本承受范围，形成“算力可得性”与“成本可负担性”的双重矛盾，成为制约其技术采纳的关键瓶颈。

技术复杂度较高则降低了中小企业的应用可行性。大模型工具的操作逻辑、维护标准与传统电商工具存在显著差异，而中小企业普遍面临专业 AI 技术人才短缺的问题。更突出的是，大模型服务商提供的本土化培训与售后指导严重不足，导致中小企业难以快速掌握技术应用要点。以亚马逊印度的 AI 工具为例，因缺乏系统性实操培训，其实际利用率不足 30%，技术应用效果较预期低 40% [2]，这一结果进一步削弱了中小企业的采纳意愿。

2.3.2. 组织维度(Organization)：能力短板与评估缺失

在组织维度(Organization)，中小企业自身的组织特征与能力局限进一步加剧了大模型技术的采纳困

境，主要表现为资源储备不足、技术与管理能力薄弱以及价值评估体系不完善三个核心问题。

资源储备不足是制约中小企业采纳大模型技术的基础障碍。一方面，中小企业资金实力有限，除了需承担高额算力成本外，还需投入数据治理、工具定制、人员培训等隐性支出，多重成本叠加导致资金压力显著增加；另一方面，中小企业的数据采集渠道狭窄、积累能力薄弱，难以获取高质量、规模化的训练数据，而缺乏优质数据支撑又会直接影响模型优化效果，最终形成“数据匮乏–模型效果差–业务提升有限”的恶性循环，进一步降低技术采纳的实际价值。

技术与管理能力薄弱则削弱了大模型技术的落地可行性。多数中小企业未组建专业的 AI 技术团队，也缺乏相关管理经验，无法独立完成大模型工具从落地部署、效果监测到迭代优化的全流程工作。例如，某跨境中小电商因不具备自主开展数据脱敏与合规审核的能力，导致大模型工具上线周期延长 3 个月；部分企业由于缺乏对模型输出结果的校验与审核能力，未能及时发现 AI 生成的虚假促销信息，直接引发用户投诉率上升，不仅未实现业务增益，反而增加了运营风险。

价值评估体系不完善则导致中小企业对大模型技术的投入持观望态度。中小企业普遍缺乏科学的大模型价值评估框架，存在三大突出痛点：其一，评估指标单一，过度聚焦成本节约等短期财务指标，而忽略用户体验提升、品牌价值沉淀等非财务指标，难以全面衡量技术价值；其二，效果归因困难，大模型的引流、转化效果与传统营销渠道、自然流量存在叠加效应，无法精准拆分其单独贡献度，导致技术价值难以量化；其三，长期价值难以捕捉，模型对用户留存率、品牌忠诚度等维度的提升效果需要长期追踪才能显现，短期内无法看到直接收益，这使得追求“短平快”回报的中小企业在投入决策上趋于谨慎。

2.3.3. 环境维度(Environment)：外部约束与生态失衡

在环境维度(Environment)，外部市场环境的复杂性与行业生态的不完善构成了中小企业采纳大模型技术的外部核心障碍，具体表现为合规监管压力不均、行业生态资源倾斜以及行业标准与支持体系缺失三个方面。

合规监管压力不均给中小企业带来了额外的运营负担与风险。不同国家及地区的数据合规法规存在显著差异，如欧盟的 GDPR、印度的《数字个人数据保护法案》等，对跨境数据传输、用户隐私保护等提出了严格要求。中小企业普遍缺乏专业的合规团队，难以全面适配各类法规条款，导致其 AI 相关业务的合规成本居高不下，且这一成本水平明显高于头部企业。合规成本的不均衡分摊，进一步压缩了中小企业的技术投入空间，成为其采纳大模型技术的重要阻碍。

行业生态资源倾斜加剧了中小企业的竞争劣势。头部电商平台与大模型服务商形成紧密合作格局，核心产业资源呈现向头部集中的趋势，算力资源、优质数据集、定制化解决方案等核心产业资源持续向头部企业集中。大模型服务商在设计产品与服务时，多以头部企业的规模化需求为核心导向，中小企业则只能被动选择功能单一、缺乏灵活性的通用版工具，难以获得匹配其差异化业务场景的技术支持。这种资源分配的失衡，使得中小企业在技术应用起点上就处于劣势，难以通过大模型技术实现竞争力提升。

行业标准与支持体系的缺失则进一步抬高了中小企业的采纳门槛。目前电商大模型领域尚未形成统一的技术标准、服务规范与价值评估准则，市场上的工具产品质量参差不齐，中小企业缺乏有效的判断依据，难以筛选出适配自身需求的优质产品，增加了技术选型的试错成本。同时，行业协会、政府部门等第三方机构提供的公共服务支持不足，如合规咨询、技术普及培训、专项资金补贴等关键服务的缺失，未能有效帮助中小企业缓解合规压力、降低技术学习成本，导致其在应对外部环境约束时缺乏必要的支撑，进一步抑制了技术采纳意愿。

2.4. 伦理层面：算法偏见凸显与内容可信性危机

2.4.1. 算法偏见加剧市场不公

大模型训练数据中的历史偏见易被放大，形成歧视性输出，违背商业公平原则。根据中国消费者协会 2023 年调查报告[11]，87.6%的受访者认为电商平台存在“区别对待不同用户群体”的现象。

在电商场景中，这种偏见主要表现为三类：一是性别偏见，例如某金融 AI 信贷模型因训练数据偏见导致女性用户通过率低于男性，类似算法偏见问题已在电商信贷审批、个性化商品推荐等场景中显现；二是区域偏见，大模型对农村地区用户的需求理解准确率显著低于城市用户，导致农村用户的推荐转化率低，加剧数字鸿沟；三是品牌偏见，头部品牌因数据积累优势，获得更多推荐曝光，可采取下列偏见指数公式来量化偏见程度。

$$\text{算法偏见指数}=\frac{\left|\text{优势群体指标}-\text{弱势群体指标}\right|}{\text{优势群体指标}}\times 100\%$$

其中，优势群体为城市用户、头部品牌等具有数据或资源优势的群体，弱势群体为农村用户、中小品牌等处于相对劣势的群体，指标为对应评估维度的量化数据(如需求理解准确率、推荐流量占比)，该指数越高说明算法偏见越严重。算法偏见不仅影响用户体验与市场公平，还可能引发舆论危机与法律风险。

2.4.2. 大模型“幻觉”引发内容可信性问题

大模型的“幻觉”特性导致其生成虚假或错误信息，严重影响电商生态信任。具体表现为三类问题：一是商品描述夸大，如 AI 生成的商品功效描述与实际不符(如宣称“美白产品 7 天见效”)，某平台因此类问题遭市场监管部门罚款 300 万元；二是虚假促销信息，如 AI 虚构折扣力度(如“原价 1000 元，限时折扣 500 元”，实际原价仅 800 元)，导致用户投诉率上升；三是跨境场景翻译错误，多语言翻译的“幻觉”问题导致商品参数翻译错误(如将“材质棉”翻译为“材质化纤”)，引发退换货纠纷与品牌信誉受损。核心成因在于：大模型的内容生成缺乏事实核查机制，部分平台为追求效率省略内容审核环节，导致虚假信息流入市场。

2.4.3. 合规备案与伦理监管体系滞后

现有监管框架难以完全适配大模型在电商领域的创新应用，企业面临“合规落地缺乏明确操作指引”的困境。国内层面，《生成式人工智能服务管理暂行办法》要求企业履行备案义务、保障内容安全，但实践中，多数电商企业因备案材料不完整、安全自评估报告不符合要求，导致备案反复被退回；行业层面，针对电商大模型的伦理准则尚未明确，对算法透明度、用户知情权的界定模糊——某平台的个性化推荐算法因未向用户披露推荐逻辑，遭用户投诉“隐私侵犯”；国际层面，不同地区的伦理监管标准差异较大，跨境电商企业需应对多重监管要求，合规成本高。此外，伦理监管的技术手段滞后，缺乏针对大模型算法偏见、内容幻觉的实时监测工具，导致部分违规行为难以被及时发现。

上述四大核心问题并非孤立存在，而是形成“技术瓶颈→数据约束→商业失效→投入不足→伦理风险→商业信任崩塌”的恶性循环。技术适配不足导致模型效果未达预期，进而加剧数据质量与合规风险；数据层面的约束进一步限制技术迭代，导致商业落地门槛高、价值难以兑现；企业因价值未达预期减少投入，使得技术优化与数据治理缺乏资源支撑；技术与数据的短板放大算法偏见与内容可信性问题，引发用户流失与卖家抵触，进一步降低商业价值，形成恶性循环。

3. 电子商务与大模型融合问题的解决路径

3.1. 技术优化：构建场景化适配体系与泛化能力提升机制

针对技术适配不足与泛化能力受限问题，需从本土化适配、冷启动突破、供应链场景优化三方面构

建协同解决方案：

第一，推进本土化语料库精准建设。联合本土企业、科研机构与行业协会，构建涵盖方言、文化习俗、消费习惯的电商专属语料库。具体建设路径包括：一是明确语料采集范围，针对印度市场，重点整合印地语、泰米尔语等方言数据及排灯节、洒红节等宗教节日消费数据；针对国内下沉市场，补充西南官话、中原官话等方言数据与低带宽场景适配数据。二是建立语料质量管控机制，采用“人工标注 + AI审核”的双重校验模式，确保语料准确性(标注准确率 ≥95%)。三是构建动态更新机制，通过实时监测用户行为数据、消费趋势报告，每月更新语料库内容，纳入新的消费场景与需求特征。

第二，优化“通用大模型 + 垂直精调”技术路径。借鉴 eCeLLM 的成功经验，通过电商领域指令微调提升模型泛化能力。针对冷启动问题，构建“少量标注数据 + 迁移学习”的优化方案：利用通用大模型的预训练优势，结合新品牌的商品属性数据、新用户的基础画像数据(少量标注，样本量 ≥500 条)进行微调，提升 unseen 场景的处理效果；开发“冷启动专用提示词模板库”，涵盖新品营销文案生成、新用户需求预判、新品推荐策略等核心场景，模板库内置电商行业术语、消费场景特征等关键信息，降低对历史数据的依赖。

第三，融合业务知识图谱与提示工程(优化后，补充具体实现逻辑与架构)。针对电商供应链、促销活动等核心场景中提示工程适配失效问题，需构建“知识结构化建模–分层注入大模型–场景化提示优化”的协同机制，将分散的电商业务知识转化为大模型可理解、可调用的结构化资源，具体实现路径如下：

(1) 电商领域知识结构化体系构建

首先对电商核心业务知识进行分类拆解与结构化建模，形成“本体库 + 规则库 + 实例库”三层知识体系，覆盖商品、供应链、促销、用户四大核心维度(见表 1)：

Table 1. Table of structured system for e-commerce core business knowledge
表 1. 电商核心业务知识结构化体系表

知识类型	核心内容	结构化形式
商品属性知识	品类分类(如 3C/美妆/生鲜)、规格参数(材质/尺寸 / 保质期)、关联关系(配件/替代品)	本体定义：商品(品类，属性{key:value}，关联商品{类型:ID})；实例数据：商品 ID:10086，品类:美妆–粉底液，属性:{质地:水润,色号:自然色,保质期:3 年}，关联商品:{替换:10087,配件:10090}
供应链规则知识	库存约束(仓库容量/安全库存阈值)、物流时效(区域–时效映射)、外部影响因素(节假日/港口政策)	规则图谱：仓库(ID，容量上限，覆盖区域) –物流时效(区域，标准时效，峰值延迟率) –约束条件(触发场景，阈值)；实例：仓库 ID:WH001，容量上限:5000 件，覆盖区域:华东，物流时效:{上海:24h,杭州:48h}，约束条件:{双十一:安全库存上浮}
促销活动知识	优惠类型(满减/折扣 / 赠品)、适用规则(品类限制/满额条件/叠加政策)、时间窗口	规则公式化：促销活动(ID，类型，适用品类，触发条件{满额:X, 数量:Y}，叠加规则，生效时间)；实例：活动 ID:Promo2025，类型:满减，适用品类:服饰，触发条件:{满 300 减 50,满 500 减 120}，叠加规则:不可与折扣券同用，生效时间:2025-11-01 至 2025-11-11
用户行为知识	消费偏好(品类优先级/价格敏感度)、场景特征(下沉市场/跨境用户)、权益等级	标签化建模：用户(ID，偏好标签{品类:权重, 价格带:区间}，场景标签，会员等级，权益范围)；实例：用户 ID:U789，偏好标签:{美妆:0.7, 3C:0.3}，价格带:[100-500]，场景标签:下沉市场–低带宽，会员等级:白银，权益范围:满减券 + 包邮

结构化过程中采用“人工梳理 + AI 辅助校验”模式：先由电商业务专家拆解核心规则(如供应链潜规则、促销叠加政策)，再通过自然语言处理工具(如依存句法分析)将非结构化规则文本转化为结构化图谱，最后通过一致性校验算法(如规则冲突检测)确保知识体系无矛盾。

(2) 结构化知识分层注入大模型机制

采用“预训练融入 + 微调对齐 + 推理调用”三层注入方式，确保知识深度融合且不丢失业务逻辑：

预训练阶段：构建知识增强预训练语料，将结构化知识转化为“知识–文本”配对数据(如商品 10086 是水润质地的粉底液，保质期 3 年，关联配件 10090)，融入大模型预训练过程，通过掩码语言模型(MLM)任务让模型学习知识与自然语言的映射关系；同时引入知识图谱嵌入(KG Embedding)技术，将实体与关系转化为向量表示，注入大模型词嵌入层，提升模型对电商实体的识别能力。

微调阶段：设计知识引导的指令微调任务，基于结构化知识生成专项微调样本(如已知仓库 WH001 容量 5000 件，华东区域物流时效 24~48 h，双十一安全库存上浮，请预测该仓库 2025 年双十一期间粉底液 10086 的备货量)，结合电商真实业务场景数据进行微调，使模型掌握“知识应用于决策”的逻辑。

推理阶段：搭建“知识图谱检索增强(KGRAG)”模块，当大模型处理具体业务请求时，实时检索结构化知识库：例如用户询问“这款粉底液能否参加双十一满减”，系统先检索商品品类(美妆–粉底液)与促销活动 Promo2025 的适用范围，将匹配结果(该商品属于服饰类满减活动适用品类，可参与满 300 减 50)作为上下文补充至提示词中，辅助模型生成准确回答(表 1)。

(3) 场景化提示工程优化策略

基于结构化知识设计“多轮引导 + 动态适配”的提示策略，避免单一提示导致的关键信息遗漏：

多轮提示分阶段引导：将复杂业务场景拆解为多轮子任务，每轮提示聚焦一个核心知识维度，逐步引导模型输出结果。以库存优化场景为例：

第 1 轮提示(基础信息输入)：已知商品 ID:10086 (粉底液，月均销量 2000 件)，仓库 WH001 (容量 5000 件，覆盖华东)，请初步预测 2025 年 11 月备货量

第 2 轮提示(供应链规则补充)：补充规则：双十一期间华东区域物流峰值延迟率 20%，安全库存需上浮 30%，仓库容量上限 5000 件，是否需要调整备货量？请说明依据

第 3 轮提示(外部因素补充)：补充信息：2025 年双十一该商品参与满 300 减 50 活动，预估销量增长 40%，竞争对手同款降价 10%，预估分流 15%销量，请最终确定备货量并给出计算逻辑

提示模板库动态适配：基于四大知识类型构建标准化提示模板库，每个模板内置知识调用占位符，根据业务场景自动填充结构化知识。例如促销推荐模板：

模板框架：用户{U789}(偏好{美妆:0.7}，价格带[100-500]，会员等级白银)当前浏览商品{10086}，当前生效促销活动{Promo2025}(满 300 减 50，适用美妆类)，请生成个性化促销推荐话术，需包含商品属性{质地:水润}与优惠规则，适配{下沉市场–低带宽}场景的简洁表达需求

提示效果反馈优化：建立“提示–输出–反馈”闭环机制，通过业务指标(如库存预测误差率、促销推荐转化率)评估提示效果，当误差率超过 10%时，自动触发提示模板优化：例如补充遗漏的知识维度(如未考虑竞争对手动态)或调整提示语序(将关键规则前置)。

(4) 系统架构图

知识结构化模块将分散的电商知识转化为标准化资源，知识注入模块通过三层机制实现知识与大模型的深度融合，提示工程模块基于结构化知识动态生成适配提示，三者协同确保大模型在电商场景中精准调用业务知识，提升输出的准确性与实用性。通过该机制，能够有效降低供应链场景下的库存预测误差，同时显著提升促销推荐的用户转化效果(图 1)。

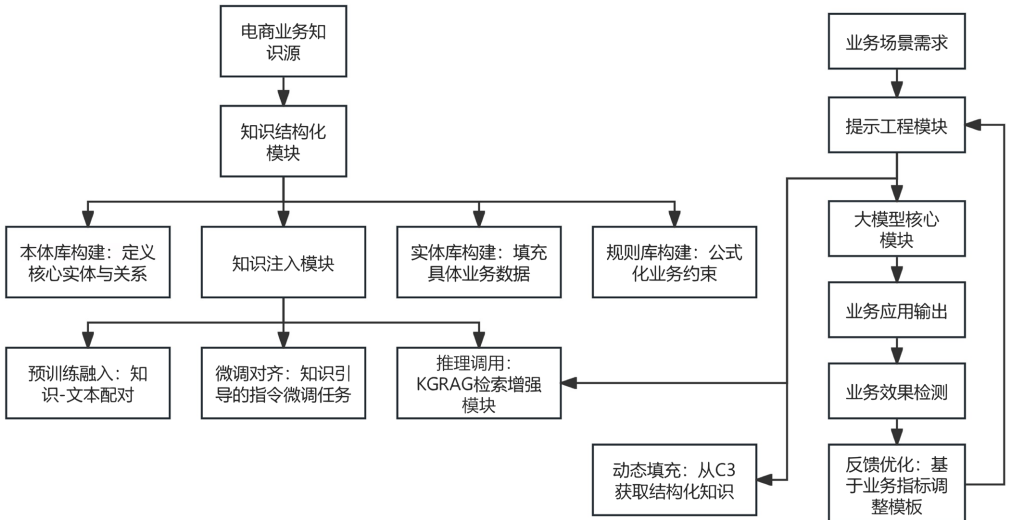

Figure 1. System architecture diagram
图 1. 系统架构图

3.2. 数据治理：建立全生命周期合规与质量管控体系

破解数据层面难题需构建“采集–处理–应用”全生命周期管理体系，强化质量管控与合规保障：

数据采集阶段：规范来源与分层授权。一是明确数据来源合法性要求，建立数据供应商资质审核机制(审核指标包括数据合规证明、数据质量评级)，禁止使用非法采集数据。二是优化用户授权协议，采用分层授权模式：核心敏感数据(如支付记录、身份信息)单独获取用户授权，明确告知数据使用范围与期限；非敏感数据(如商品浏览记录)采用默认授权 + 可撤回模式，保障用户知情权与选择权。三是搭建行业级数据共享平台，采用联邦学习技术实现“数据可用不可见”，平台由行业协会主导建设，制定统一的数据共享标准与收益分配机制，缓解中小企业数据短缺困境。

数据处理阶段：强化质量管控与脱敏处理。一是建立多源数据清洗与质量评估机制，采用 AI 算法(如聚类算法)自动识别刷单数据、无效评论等噪声数据，通过插值法、生成式 AI (如 GAN 模型)填补缺失数据；建立数据质量考核指标体系(标注准确率 ≥95%、噪声率 ≤5%、数据完整性 ≥90%)，定期开展数据质量审计(每季度 1 次)。二是推广差分隐私、数据脱敏等技术，对用户支付记录、收货地址等敏感信息进行加密处理：采用动态脱敏技术，对不同权限人员展示不同粒度的数据(如普通员工无法查看完整地址)；通过差分隐私技术在数据中加入微小噪声，避免敏感信息泄露，同时确保脱敏后数据仍保持可用性。

数据应用阶段：强化合规与安全保障。一是严格遵守数据本地化存储要求，针对跨境电商企业，构建跨境数据传输合规通道：通过 GDPR 认证的跨境数据传输机制(如标准合同认证)，或在目标市场建立本地数据中心，确保数据传输合规。二是建立数据安全实时监测系统，部署异常访问检测、数据泄露预警等功能，对敏感数据的访问、传输进行全程日志记录，发现异常行为立即触发预警(响应时间 ≤ 10 分钟)。三是推出中小企业数据治理 SaaS 服务，由第三方服务商提供低成本的数据清洗、脱敏、合规审核工具，降低中小企业技术门槛与成本(服务费控制在年均 5 万元以内)。

3.3. 商业赋能：降低落地门槛与构建科学价值评估体系

针对商业落地困境，需通过轻量化工具、场景化适配、价值评估体系构建三方面协同发力：

第一，搭建轻量化 SaaS 服务平台。由头部企业或第三方服务商提供低成本、易操作的大模型工具包，涵盖三大核心功能模块：智能运营模块(商品标题优化、营销文案生成、客服自动回复)、简易数据分析模

块(销售趋势预测、用户画像分析)、低带宽 AI 客服模块。平台推出免费基础版与付费进阶版，免费版覆盖中小卖家核心需求，付费进阶版(年均服务费 ≤3 万元)提供定制化功能。同时，提供本土化培训与售后指导：通过线上课程(如短视频教程、直播培训)、线下 workshops 等形式，提升中小企业操作能力；建立 7 × 12 小时售后客服团队，及时响应技术问题。例如，拼多多“AI 商家助手”针对中小企业推出免费基础版，包含商品标题优化、订单自动处理等功能，覆盖超 800 万中小卖家，工具利用率达 65%。

第二，推进“技术–业务”深度融合。基于电商业务痛点设计场景化解决方案，避免技术套用。具体措施包括：针对农村卖家推出低带宽适配的 AI 工具，优化加载速度(离线模式下加载时间 ≤3 秒)与离线使用功能；针对跨境电商翻译场景，开发“AI 翻译 + 人工复核”机制，建立翻译错误反馈数据库，持续优化翻译模型，将翻译误差率控制在 5%以内；建立技术应用效果反馈机制，通过卖家调研、用户评价、业务数据监测等方式，实时收集意见，每季度迭代优化工具功能。

第三，构建多维度价值评估体系。建立“技术投入–业务指标”映射模型，涵盖三大核心维度：① 财务指标(成本节约率、收入提升率)，计算方法为成本节约率 =(传统模式成本 −AI 模式成本)/传统模式成本 ×100%，收入提升率 =(AI 模式收入 − 传统模式收入)/传统模式收入 ×100%；② 非财务指标(用户体验提升率、品牌价值增长率、合规风险降低率)，其中用户体验提升率可通过 NPS 评分变化、投诉率下降幅度衡量；③ 长期价值指标(用户留存率、品牌忠诚度提升)，通过 6 个月以上的纵向追踪数据评估。采用多触点归因模型(如马尔可夫链模型)拆分大模型与其他渠道的贡献，明确大模型的实际价值，为企业投入决策提供依据。

3.4. 伦理监管：建立算法治理框架与合规保障机制

化解伦理合规风险需构建多元协同治理体系，结合全生命周期合规流程：

政府层面：完善监管细则与技术赋能。一是加快出台电商大模型专项监管细则，明确算法备案流程(如简化中小企业备案材料)、内容审核标准(如 AI 生成内容的事实核查要求)与伦理准则(如禁止算法歧视、保障内容可信)。二是建立分级分类监管机制，针对不同规模企业、不同风险等级应用实施差异化监管：头部企业重点监管算法偏见、数据隐私保护；中小企业重点提供合规指导，降低监管成本。三是搭建监管技术平台，利用 AI 工具实时监测算法偏见、虚假信息等违规行为，提升监管效率[4]。

企业层面：强化自律与内控机制。一是建立算法透明度机制，向用户披露推荐、定价等算法的基本原理与数据来源(如通过 APP 隐私政策、算法说明页面展示)；对高风险算法(如信贷评估、个性化定价)开展公平性评估，定期发布算法伦理报告。二是设立 AI 伦理审查委员会，由技术、法律、业务等多领域专家组成，负责算法偏见监测、生成内容审核等工作，建立“生成–审核–发布”三级内容管控流程：AI 生成内容先由算法初步审核，再经人工复核，最后发布。三是完善合规备案流程，按照《生成式人工智能服务管理暂行办法》要求准备材料，开展安全自评估，聘请第三方机构进行合规认证，确保备案顺利通过。

行业层面：推动共治与标准制定。一是成立电商 AI 联盟，由行业协会、头部企业、科研机构组成，制定行业伦理规范与技术标准(如数据脱敏标准、算法偏见评估指标、AI 生成内容审核标准)。二是开展合规培训与能力建设，为中小企业提供免费合规咨询服务，组织合规案例分享会，提升行业整体合规水平。三是建立行业黑名单制度，对存在严重伦理违规(如故意生成虚假信息、算法歧视)的企业进行惩戒，限制其市场准入，引导行业规范发展。

4. 总结与展望

4.1. 研究结论

电子商务与大模型的融合是数字经济发展的必然趋势，但目前仍面临四大核心问题(按影响优先级排

序)：技术层面的场景适配与泛化能力不足(首要瓶颈)、数据层面的质量管控与合规风险凸显(基础约束)、商业层面的中小企业落地门槛与价值兑现困境(关键障碍)、伦理层面的算法偏见与内容可信性危机(生态保障缺失)。这些问题相互交织形成恶性循环：技术适配不足导致模型效果未达预期，加剧数据质量与合规风险；数据约束进一步限制技术迭代，推高商业落地门槛；企业因价值未兑现减少投入，导致技术优化与数据治理缺乏资源；技术与数据短板放大伦理风险，引发商业信任崩塌，反向降低商业价值。

解决上述问题需从技术、数据、商业、伦理四个维度协同发力：技术层面构建场景化适配体系与泛化能力提升机制；数据层面建立全生命周期合规与质量管控体系；商业层面降低中小企业落地门槛并构建科学价值评估体系；伦理层面建立“政府监管－企业自律－行业共治”的算法治理框架。四大路径相互支撑，共同破解恶性循环，推动融合发展从“技术赋能”向“生态重构”转型。

4.2. 未来展望

未来，随着大模型技术的持续迭代与行业生态的不断完善，电子商务与大模型的融合将呈现三大发展趋势：

技术层面：多模态融合与场景深度渗透：多模态大模型与数字孪生技术的结合，将实现电商“人货场”的全场景虚拟重构，提升用户沉浸式体验；Agent 技术在电商供应链的应用，将实现需求预测、库存优化、物流调度的全流程自动化；低代码、无代码大模型工具的普及，将进一步降低技术应用门槛。

商业层面：中小企业普及与价值精准化：轻量化 SaaS 工具的持续迭代将推动大模型在中小企业的普及应用，行业数字化鸿沟逐步缩小；价值评估体系的完善将实现 ROI 精准量化，引导企业理性投入；“技术－业务”深度融合将催生更多创新应用场景(如 AI 驱动的个性化定制、智能跨境贸易全流程服务)。

监管层面：技术治理与制度监管协同：“技术治理 + 制度监管”的协同模式将趋于成熟，AI 监管工具的应用将提升监管效率；行业伦理规范与技术标准的完善将引导企业合规发展；跨境合规协作机制的建立将缓解跨境电商企业的合规压力。

本研究虽系统识别了融合发展的核心问题并提出解决路径，但仍存在一定局限性：未对不同规模、不同类型电商企业(如综合电商、垂直电商、跨境电商)的问题差异进行细分研究，未来可结合实证调研，开展针对性研究；同时，随着技术与政策的动态变化，融合发展中可能出现新的问题(如多模态大模型的伦理风险、AI 生成内容的知识产权问题)，需持续追踪并完善解决方案。

参考文献

[1] 罗长银, 陈学斌, 张淑芬. 基于联邦集成算法对不同脱敏数据的研究[J]. 应用科学学报, 2024, 42(4): 94-102.

[2] 澎湃新闻. 在印度, 亚马逊的 AI 连“生日蛋糕”都听不懂[EB/OL]. 澎湃新闻. https://www.thepaper.cn/newsDetail_forward_32157573, 2025-12-12.

[3] 翟正同, 尹浩华. 人工智能技术在跨境电商领域的应用探索[J]. 电子商务评论, 2025, 14(1): 3910-3918. https://doi.org/10.12677/ecl.2025.141485

[4] 中国电子技术标准化研究院. 生成式人工智能数据应用合规指南[EB/OL]. http://cecc-iempc.cn/newsinfo/7139479.html, 2024-04-16.

[5] 吴道生. 大模型驱动的智能客服在电子商务中的应用与挑战[J]. 电子商务评论, 2025, 14(11): 134-141.

[6] 高振. 语言大模型在电子商务中的应用[J]. 中国电子商务, 2024(8): 33-36.

[7] Ning, L., Zhang, H., Liu, Y., *et al.* (2025) eCeLLM: Generalizing Large Language Models for E-Commerce from Large-scale, High-Quality Instruction Data. *IEEE Transactions on Knowledge and Data Engineering*, **37**, 1890-1903.

[8] 张卫东, 陈希鹏, 李松涛. 多维框架下个性化推荐系统的流行度偏见测度方法与实证研究[J]. 情报资料工作, 2024, 45(2): 66-74.

[9] 陈晨, 赵宇. 电商大模型算法偏见的形成机制与规制路径[J]. 情报杂志, 2024, 43(5): 167-173.

[10] 跨境电商 AI 智能体解决方案: 数商云的技术架构与全链路赋能[EB/OL]. https://m.shushangyun.com/article-28585.html, 2025-08-15.

[11] 中国消费者协会. 中国消费者权益保护状况年度报告(2023) [N]. 中国消费者报, 2024-05-27(01).