• 专题:双清论坛“大模型时代数智商务的理论与方法” •
DOI: 10.3724/BNSFC-2025.06.07.0002

大模型时代数智商务:理论进展、关键科学问题与方法路径*

徐 心[1] 叶 强[2] 王汉生[3] 任之光[4] 姜广鑫[5] 高超越[2]
夏 昊[5] 解海天[3] 胡丹琪[3] 常嘉路[1] 林志杰[1**]

1. 清华大学 经济管理学院,北京 100084
2. 中国科学技术大学 管理学院,合肥 230026
3. 北京大学 光华管理学院,北京 100871
4. 国家自然科学基金委员会 管理科学部,北京 100085
5. 哈尔滨工业大学 经济与管理学院,哈尔滨 150001

[**摘 要**] 面向国家重大需求与学术研究前沿,大模型时代数智商务的发展不仅依赖于人工智能等关键技术的突破及领域应用,也要求传统商务流程再造与管理理论创新。深度融合数据智能与商务场景,突破核心技术瓶颈,催生新产业与新模式,实现智能、高效、协同的数智商务新格局是未来重要的发展方向。基于国家自然科学基金委员会第406期“双清论坛”,本文围绕大模型时代数智商务的理论和方法,聚焦数智商务的“关键技术与方法”、“智能协同与自主决策”以及“新模式与新业态”三方面进行了深入探讨,并进一步提出针对人工智能方法与大模型技术突破、多智能体决策与跨智能体协同、人智融合的流程—组织—市场再造的研究议程,及优先支持面向数智商务的管理科学与计算科学交叉领域研究的资助建议。

[**关键词**] 大模型;数智商务;多智能体;人智协同;治理体系;组织变革

随着信息技术,特别是互联网、物联网、云计算、人工智能等新技术的蓬勃发展,以大模型为代表的通用人工智能技术正推动数智商务进入范式变革的新阶段。大模型凭借其强大的语义理解、知识涌现和复杂推理能力,正在重构传统商务活动的决策逻辑与价值创造模式。在金融科技领域,基于Transformer架构的智能投顾系统已实现跨市场、多模态数据的实时决策[1];在供应链管理中,多智能体强化学习框架显著提升了复杂网络下的协同优化效率[2,3];而在消费者行为分析方面,大模型驱动的动态画像技术正突破传统计量模型的局限[4],实现微观个体需求与宏观市场趋势的精准对接。这些实践创新不仅验证了数智技术在商务场景中的变革潜力,更揭示了智能技术与商业规律深度融合所催生的基础理论挑战。

党的二十大报告强调加快发展数字经济,推动数实融合的经济发展新格局。随后,二十届三中全会通过的《中共中央关于进一步全面深化改革 推进中国式现代化的决定》(以下简称《决定》),在党和国家的政策文献中首次引入“数智”概念,这一表述深刻体现了数字化与智能化技术融合对于当前发展的关键作用,并明确指出数智商务是驱动数字经济快速发展、促进数字经济与实体经济深度融合的核心动能。为响应国家战略需求,国家自然科学基金委员会第406期双清论坛“大模型时代数智商务的理论和方法”于2025年4月在北京召开,汇聚管理科学、计算机科学、经济学、统计学等多领域学者,重点研讨了大模型赋能下的商务智能关键技术与决策范式、人智融合的新型商务组织与模式以及数智商务变革背

收稿日期:2025-06-07; 修回日期:2025-10-27

* 本文根据国家自然科学基金委员会第406期“双清论坛”讨论的内容整理。

** 通信作者,Email:linzhj@sem.tsinghua.edu.cn

引用格式: 徐心,叶强,王汉生,等. 大模型时代数智商务:理论进展、关键科学问题与方法路径. 中国科学基金,2025,39(5):686–704.
Xu X,Ye Q,Wang HS,et al. Digital intelligence business in the era of large language models:Theoretical advances,key scientific issues,and methodological approaches. Bulletin of National Natural Science Foundation of China,2025,39(5):686–704.(in Chinese)

景下的模式创新与管理等前沿议题。论坛指出,亟需建立适配新型前沿技术的数智商务理论框架,突破数据—算法—场景协同创新的方法论瓶颈,这对培育新质生产力、构建中国自主的知识体系具有双重战略意义。当前,数智商务研究正呈现出“技术驱动—问题导向—价值引领”的三元特征,其理论发展不仅需要解决大模型垂域化落地等基础科学问题,更需构建符合数字经济伦理的智能治理体系,为全球数字经济发展贡献中国方案。

1　大模型时代数智商务的概念

本次论坛关注的核心是大模型时代的数智商务。数智商务以人工智能技术和数据要素为基础,通过人机协作和融合,完成自主决策过程,构建自动化、智能化、协同化、网络化、集成化的产业模式,成为经济活动价值创造的核心支撑。

伴随大模型、智能体等人工智能前沿技术的加速突破,数智商务正在从众多产业场景中展现出广阔的发展前景,涵盖数智营销、数智供应链、组织管理创新与金融科技等核心领域。在数智营销领域,不同市场主体之间的边界变得模糊,并逐渐成为价值共创的合作者[5]。在数字技术的支持下,供应方和需求方实时互动增强,随着用户需求的动态变化,产品和服务快速迭代升级,营销策略也更加动态化、多样化、个性化[6]。数智技术也影响到供应链的每个环节,成为供应链生态系统形成的重要支撑[7]。特别是人工神经网络的应用,通过对复杂模式的分类、聚类、预测、检索等过程控制,有效提高了采购、生产、仓储、运输、销售等各个环节的协同效率[8]。以物联网基础设施为支撑的“零库存智能制造”,成了数智供应链的新型制造范式[9]。区块链技术则赋能供应链各个环节的信息可视化、合同自动化,增强信息韧性,促进供应链与金融一体化发展。

大模型时代的数智商务将前沿技术与生产经济紧密联系。供给端要求创新组织方式、流程整合和要素协同以匹配技术、业务与生态协同演化的新规则;需求端则关注用户对智能技术的认知与情感的技术效能影响。同时,需探索算法治理、契约治理和关系治理的混合治理机制,从而提高供需双方的匹配度和融合度[10]。数智商务的发展需要管理、数理、信息、心理等多学科研究范式的更新,以突破常规科学解密失效,为数智商务发展提供持续的动力和探索空间。

2　大模型时代数智商务的理论和方法的主要研究内容及进展

大模型时代数智商务有着丰富的落地场景以及深厚的理论研究基础。对相关的研究内容以及进展,将从三方面展开综述:(1)面向数智商务的关键技术与方法;(2)数智商务系统的智能协同与自主决策;(3)数智商务变革的新模式与新业态。图1描述了这三个方面之间的关系,内容各有侧重,且交叉融合、互相促进。

2.1　面向数智商务的关键技术与方法

在数智商务时代,商务垂直行业场景呈现出高度复杂、多样与动态演化的特征,人工智能技术,特别是大模型技术的泛化应用面临巨大挑战。面对多源异构数据、多变业务需求及高风险运行环境,亟需将人工智能技术进行深度定制,并融合行业知识以增强模型对业务语境的理解与适应能力。在实际应用中,模型部署不仅受限于算力资源与数据可用性,还需满足合规、安全与隐私等多重要求。同时,数智商务的发展对智能技术提出了更高的实时性、稳健性与可解释性要求,传统的通用建模与分析方法已难以满足当前复杂业务场景下的智能决策需求。研究面向数智商务的关键技术与方法,不仅是推动人工智能从实验室走向产业落地的关键路径,也是支撑企业数字化转型、提升业务智能水平的核心基础。

面对上述挑战,未来的研究将围绕“模型部署—领域增强—风险控制”三阶段研究路径,提出一系列核心技术与方法。首先,在模型部署方面,重点研究在资源受限的真实业务环境中部署并高效适配人工智能模型,解决轻量化建模与迁移学习等问题。在领域增强方面,借助强化学习与领域知识融合,研究模型性能与业务效益的协同提升问题。在风险控制方面,研究以风险控制为核心,建立稳健、可解释、合规的数据与算法框架,以保障模型在复杂商业环境中的安全可控运行。具体的研究涉及以下五个方面。

2.1.1　资源受限商务场景下人工智能算法研究

在数智商务实践中,企业普遍面临算力资源有限和部署环境复杂的现实问题,开发适用于资源受限场景的人工智能算法成为研究的重点。针对这一问题,学者们提出了多种优化和改进方法,主要集中在大模型的高效微调与推理优化,推理阶段的参数压缩技术,小模型的结构设计与应用,以及知识蒸馏与迁移学习的改进这四个方面。图2展示了资源受限商务场景下的主要工程策略。

在资源受限场景下,大模型的高效微调与推理优化是一个重要的研究方向。通过高效的微调技术,可以降低模型训练和部署所需的计算资源,同时实现模型能力的快速适配。目前,低秩分解、参数冻结、增量学习等方法被广泛应用于这一领域。如Hu等[11]提出的低秩分解

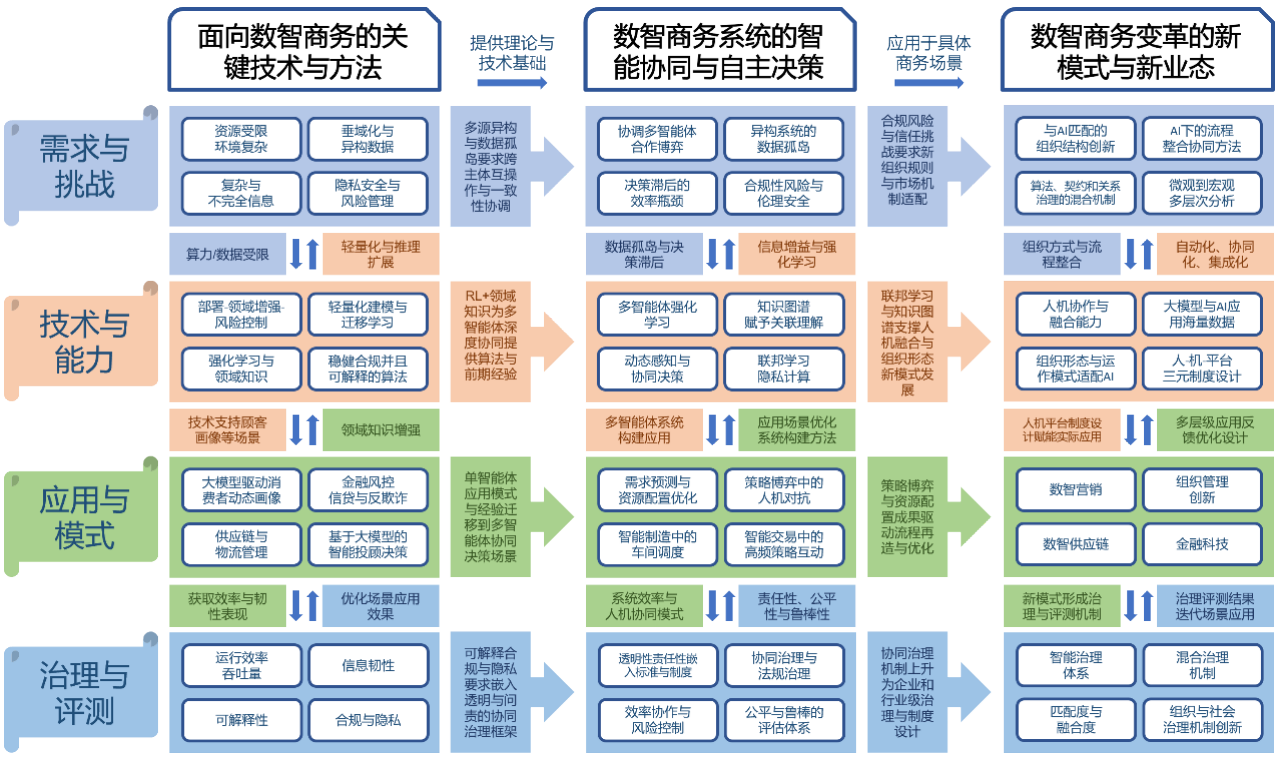


图1 大模型时代数智商务的理论与方法——总体概念模型

Fig.1 Theory and Methods of Digital Intelligence Business Conceptual Model in the Era of Large Language Models

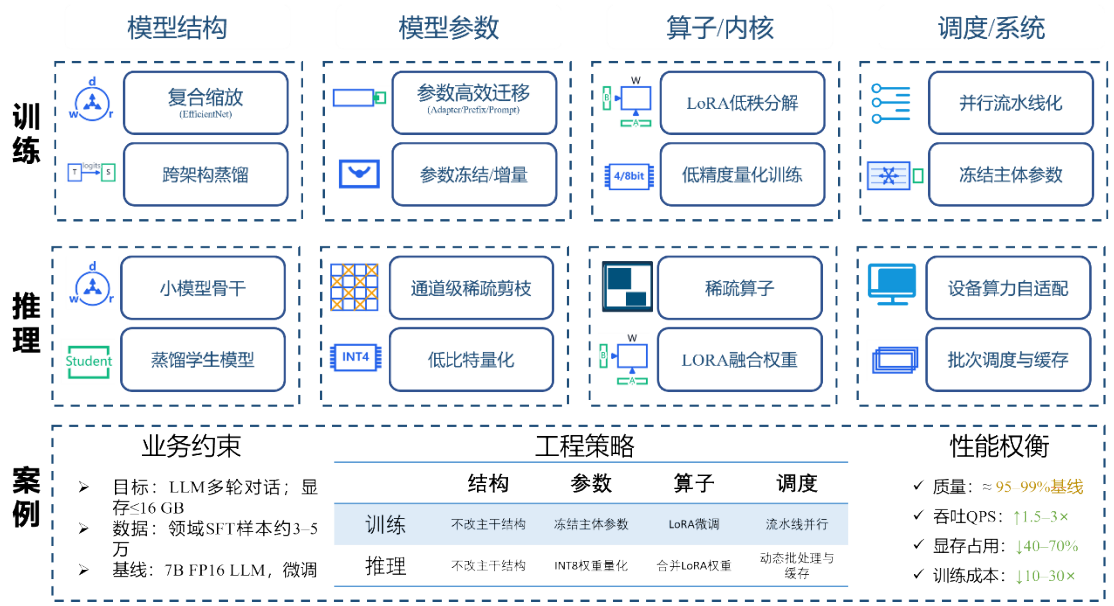


图2 资源受限商务场景下的工程策略

Fig.2 Engineering Strategy in Resource Constrained Business Scenarios

方法，通过冻结预训练模型的权重，将可训练的低秩分解矩阵引入到Transformer架构的各层中，大幅减少了微调时的参数数量和显存需求。与其他微调方法（如适配器层）相比，低秩分解在推理阶段没有额外的计算延迟，并显著降低了硬件门槛，使得在资源受限场景中部署大模型成为可能。另外，Saha等[12]提出了随机低秩与低精度分解方法，结合低秩分解和量化，通过随机投影近似矩阵的列空间，对结果进行低精度量化，显著减少了存储和计算成本。

推理阶段的参数压缩是解决资源受限问题的另一重要路径。通过剪枝、量化和知识蒸馏等技术，可以在不显著降低模型性能的前提下，减少模型的存储和计算

成本,从而适应低算力设备的需求。如Sun等[13]提出了一种稀疏化剪枝方法优化大模型的参数压缩效率。该方法通过引入权重和输入激活的乘积作为剪枝指标,在不需要权重更新的情况下实现了高效的剪枝过程。与传统的全局剪枝方法不同,该方法基于输出通道逐行比较权重的重要性,从而确保在高稀疏度下依然能够保持模型性能。

除了对大模型进行优化,小模型的设计、训练与应用也是资源受限场景中的重要研究方向。相比大模型,小模型在计算资源需求上更具优势,但如何在轻量化设计的同时保证模型性能,是一个关键问题。Tan和Le[14]提出的Efficient Net是小模型设计的一个典型代表,其核心思想是通过复合缩放方法,在网络的深度、宽度和分辨率三个维度上进行均衡扩展,从而最大化模型的性能与效率。具体而言,复合缩放方法通过引入一个复合系数,将资源分配到网络的三个缩放维度,确保模型在扩展时保持结构的平衡性。这种方法不仅避免了单一维度扩展带来的性能瓶颈,还能够通过一次小规模的搜索确定缩放系数,从而显著降低设计复杂度与计算成本。

知识蒸馏与迁移学习为资源受限场景下的大模型向小模型的知识转移提供了重要途径。通过知识蒸馏技术,可以将预训练大模型中蕴含的丰富知识压缩并迁移到小模型中,从而显著提升小模型的表现能力。在这方面,Hao等[15]提出了一个通用的跨异构架构知识蒸馏框架,该框架针对现有蒸馏方法在异构架构间表现不佳的问题,通过引入额外的出口分支将中间特征投影到对齐的logits空间,从而消除架构特定信息的干扰,实现了CNN、Transformer和MLP等异构模型之间的高效蒸馏。另外,He等[16]梳理了现有方法(如Adapter、Prefix Tuning、Prompt Tuning等)的设计共性,提出了一种迁移学习框架,在冻结大模型主体参数的基础上,仅通过修改特定隐藏状态以实现知识迁移。通过组合已有设计元素构建出新的微调方法,该框架显著降低了迁移过程中的参数开销,同时在多任务下保持了与全参数微调相当的性能,验证了其在低资源环境中实现高效知识迁移的实用价值,展示了高效迁移策略在多任务场景下的广泛适用性与优越性能。

2.1.2 *面向数智商务的大模型垂域化关键技术研究*

垂直商务场景通常涉及复杂的业务逻辑、多样化的数据形式以及高度动态的环境需求,现有通用模型在适应这些特定场景时面临诸多挑战,需要在精准性、实时性和业务适配性上做出显著提升,这对模型的构建和应用提出了更高要求。为此,亟需结合实际业务特征,从大模型的开发、适配到应用全流程开展针对性研究。

垂直商务场景的数据形式通常呈现出多源异构的特点,包括结构化数据(如财务报表、库存记录)、非结构化数据(如客户评论、图像、视频)以及实时流数据(如交易日志、物联网传感器数据)。如何有效整合并治理这些多模态数据,构建大模型多模态数据开发与治理框架,是提升大模型认知能力的首要挑战。要实现多模态数据的高效整合与治理,必须充分挖掘不同模态数据之间的互补信息,同时确保数据的高质量和一致性。在这个方面,Li等[17]提出了一种新型多模态融合模型,该模型采用了一种耦合状态转移机制,将每个模态的当前状态与其他模态的历史状态相结合,从而实现对多模态特征的深度融合。同时,该模型还针对硬件并行化设计了全局卷积核,显著提高了计算效率。

大模型训练和推理通常需要消耗大量算力资源,而算力资源的高效调度与分配是保证模型性能和成本可控的关键。要实现性能与成本的平衡,需要设计动态调度机制,同时优化计算资源的利用率。如Mirhoseini等[18]提出了一种基于深度强化学习的图优化方法,该方法将芯片布局设计问题抽象为一个强化学习任务,使用图神经网络对芯片结构进行建模,并通过强化学习代理逐步优化资源的分配。与传统的分区方法和解析求解器相比,该方法能够直接优化可微和不可微的目标函数,例如功耗、延迟和布线拥塞等复杂指标。此外,该方法通过学习可迁移的表示,能够在不同的任务间实现快速泛化,从而显著提高新任务的优化效率。

在垂直商务场景中,大模型需要在通用知识迁移和业务特征刻画之间找到平衡。传统的端到端模型设计通常难以同时满足这两方面的需求,而模块化的协同架构通过耦合与解耦不同模块的功能,为这一问题提供了有效的解决方案。通过不同任务模块设计,可以在统一框架下实现通用能力的迁移与特定需求的优化,从而提升模型的适配效率和推理能力。如Zhang等[19]提出了一种新的预训练方法,专门用于提升抽象式文本摘要的性能。其核心创新在于设计一个“间隙句子生成”的自监督预训练目标,即从文档中移除或掩码若干重要句子,并训练模型根据剩余内容将这些被移除的句子生成出来,从而使预训练过程更贴近摘要任务。

在动态变化的垂直商务场景中,智能体的全生命周期管理是确保其长期高效运行的核心。全生命周期管理包括从数据采集、模型训练、部署到持续优化的各个阶段,核心目标是建立从数据采集、模型训练到部署优化的闭环系统,以保障模型在复杂环境下的稳定性和持续改进能力。Zheng等[20]通过引入内部知识和外部知识两类持续学习机制,系统性地探讨了模型在不同生命

周期阶段的适应性问题。内部知识方法通过持续预训练和微调,将新知识直接融入模型参数;外部知识方法则通过检索增强生成和工具集成等手段,动态扩展模型的能力,而无需更新其核心参数。这种分类方法为智能体的全生命周期管理提供了全面的视角。

2.1.3 *复杂商务决策场景下的大模型强化学习*

复杂商务决策环境具有高维、多主体交互、动态变化与部分可观测等核心特征。一方面,现代商业系统中的政府、企业、消费者等多元主体因目标差异展开持续博弈,各主体协同创新与策略冲突并行,海量要素交叉作用形成高维参数空间,导致网络的复杂性在数字化浪潮中呈指数级放大;另一方面,商业环境持续受技术突破、地缘政治、社会趋势等外生变量冲击,通过主体行为形成不断动态变化的商业网络,导致网络中的关键信息获取具有强滞后性,即便通过大数据监测,仍面临数据噪声、语义断层等干扰,迫使决策者始终在信息不完全中行动,为决策模式带来巨大挑战。因此,应以高效计算与动态可扩展性为核心,结合统计计算与大语言模型等技术,突破传统分析与优化方法的固有局限,构建适配高维动态博弈的智能分析与决策范式。

由于商务环境难以大规模实验,需整合多源数据构建商务决策场景的仿真环境模型,解决环境随机性、多主体博弈与长周期决策建模问题。此类建模需融合微观主体行为与宏观系统演化,通过动态博弈理论、随机过程及机器学习方法量化不确定性因素。如Huang等[21]针对直播促销场景,采用几何布朗运动建模直播间实时人气的随机波动,并将促销策略优化转化为脉冲控制问题,通过仿真验证了策略对多主体竞争环境的适应性,为长周期决策提供了动态调整框架。此类方法论可迁移至新兴市场商务决策建模,通过数据驱动的仿真系统捕捉政策、文化及市场动态的交互影响,从而提升决策的鲁棒性与前瞻性,为复杂商务决策提供可解释的理论框架与实践指导。

在复杂商务决策场景(如供应链动态优化、多产品库存管理、实时定价策略)中,高维状态—动作空间与稀疏奖励信号导致传统强化学习方法面临探索效率低、样本成本高、策略泛化性差等挑战。大规模高精度强化学习方法是解决此问题的有效方法。特别地,分层强化学习(Hierarchical Reinforcement Learning, HRL)与多任务迁移学习协同框架可以通过分层抽象与知识共享机制,提升大规模高精度强化学习在商务场景中的决策效率与鲁棒性。如Yan等[22]为解决多产品拆零销售场景下的组合订购决策问题,提出了一种融合分布鲁棒优化(Distributionally Robust Optimization, DRO)与深度强化学习(Deep Reinforcement Learning, DRL)的混合框架。该研究将产品需求的不确定性建模为均值—协方差约束的模糊集,通过DRO生成鲁棒订购量的先验知识,并设计策略迁移机制将其注入DRL的探索过程。当前在分层策略的信用分配难题以及跨任务表征的泛化边界问题等方面仍存在一定的空白。如何捕获商务决策中的长程依赖关系,以实现更高效的在线策略迭代仍是需要处理的难题。

复杂商务决策场景常面临多目标冲突、奖励稀疏性及环境扰动等导致传统强化学习面临策略失效风险的困境,因而有必要设计复杂环境下奖励通用框架。多目标冲突下的信用分配、基于人类反馈的奖励建模及对抗性扰动下的鲁棒性问题将成为复杂商务决策场景研究的重要前沿。如Chen等[23]提出基于影响矩阵与距离矩阵的动态信用分配模型。通过量化智能体间的空间关联与行为影响,构建可解释的奖励系数矩阵,将全局商务目标分解为局部可优化子任务;Dubois等[24]基于大语言模型生成人类反馈的仿真数据用于训练奖励模型,验证了仿真反馈与真实人类反馈的高度一致性,为低成本研究基于人类反馈的强化学习提供了有效框架。

除了训练方法与奖励框架设计问题,强化学习智能体还面临仿真环境与真实场景的差异、在线部署时的策略失效及安全约束缺失等挑战。强化学习智能体的虚实环境迁移、领域自适应、在线策略微调及安全约束部署对于提升AI智能体在实际应用中的稳定性与可靠性有重要意义。如Li[25]提出了合作式上下文学习,建立起了跨区域和跨领域的长程依赖关系,从而促进了语境知识的转移,并简化了适应过程;Yu等[26]提出了一种新的激活修剪方案,通过稀疏化的反向传播策略,巧妙地避免了不必要的激活修剪、存储和梯度计算,显著提高了在内存受限设备上微调基础模型的效率。

2.1.4 *领域增强的商务智能技术与方法*

在数智化转型加速的背景下,通用大模型在商务场景中的应用通常面临领域知识缺失导致决策依据不足,逻辑严谨性不足引发推理偏差,专业术语理解偏差降低结果可信度等挑战。领域增强技术通过融合行业知识、理论框架与垂直数据,构建兼具泛化性与专业性的智能决策系统,其核心在于实现领域知识与数据驱动范式的深度耦合,破解“黑箱”导致的决策风险与可解释性困境。

在领域知识增强方面,通用大模型因缺乏对垂直领域规则、流程及隐性经验的深度理解,难以满足复杂商务场景的精准决策需求。领域知识增强技术通过结构化嵌入行业规则、专家经验等知识,成为破解这一困境

的核心路径。如Liu等[27]提出了一种新型大语言模型(Large Language Models, LLMs)蒸馏框架,根据教师和学生模型在领域性能上的差异,动态地、平滑地调整知识蒸馏数据集的组成,使知识蒸馏过程更加稳定和有效,显著提高了学生模型性能。

在领域理论增强方面,商务决策依赖博弈论、供应链网络理论等学科范式,需将其抽象为可计算的约束条件或损失函数。例如供应链优化中可引入纳什均衡约束,金融风控中嵌入Black-Scholes方程,通过嵌入物理信息的神经网络实现理论框架与数据驱动的协同优化。在数智化转型背景下,通用大模型因领域知识缺失、逻辑严谨性不足及专业术语理解偏差,难以满足复杂商务场景的精准决策需求。领域模型增强技术通过构建轻量化专业模型、优化领域适配预训练机制及模块化架构设计,成为破解这一困境的核心路径。针对垂直场景的碎片化需求,需训练轻量化专业模型,研究领域适配的预训练(如用供应链语料微调大模型)、小样本学习(如新兴市场预测)、模型模块化(如插件式风险控制组件)等技术。如Wu等[28]提出将经典深度网络技术迁移到元级的策略以进一步改进上下文学习(In-Context Learning,ICL),提升了少量数据领域微调快速适应能力。

2.1.5 商务智能领域的安全与风险管理关键技术

商务智能技术已深度应用于金融交易、医疗健康、数字政务及工业数字孪生等关键领域,其处理的核心数据兼具双重属性:既包含用户隐私、商业机密等高价值敏感信息,也涵盖气候监测、城市治理等公共数据资产。这类数据普遍呈现多源异构、时空关联性强、动态演化显著等特征,在采集、传输与共享过程中面临三重核心风险:原始数据层面的隐私泄露与恶意篡改风险、数据处理环节的模型偏差风险,以及决策应用中的合规失效风险[29,30]。在实践中,差分隐私的隐私预算 (ε)是衡量隐私保护强度的核心指标,其数值越小,隐私保护程度越高,但可能伴随数据效用损失。研究实践中,应根据具体的应用场景和隐私需求,在数据效用与隐私强度之间进行权衡,选取合适的 ε预算。安全多方计算则允许多方在不暴露原始数据的情况下协同计算约定函数,其安全性依赖于密码学协议。尽管这些技术能够有效解决现在大多数隐私保护问题,然而在金融与医疗等高敏场景下隐私保护程度和计算开销之间的均衡仍存在一定的研究空白,如何针对高敏商务智能场景设计有效的隐私保护算法是目前商务智能风险管理领域的一大难题。

联邦学习(Federated Learning,FL)是一种分布式机器学习范式,旨在保护数据隐私的前提下,实现多方数据联合建模。其核心目标是解决数据孤岛问题,通过分布式训练和隐私保护技术,使多个参与方在不共享原始数据的情况下协同训练模型。Zhou等[31]基于联邦学习有效地解决基因组测序的隐私保护问题。但如何在通信效率、异构数据融合与恶意攻击防御之间建立动态平衡机制是目前商务智能的一大挑战。具体而言,联邦学习框架应关注并优化以下关键指标:在通信效率方面,应量化模型更新的传输数据量与原始模型参数量的比值,即压缩比,并通过梯度量化、稀疏化或差分编码等技术进行优化,以适应带宽受限的实际部署环境;在异构数据融合方面,应评估模型在具有不同数据分布的多个客户端上训练后,中心模型在未知数据上的泛化性能,确保其性能与在理想同构数据上训练的模型性能保持可比性或达到预设的最低性能标准;在恶意攻击防御方面,需量化系统对成员推断攻击(Membership Inference Attack,MIA)和数据投毒攻击(Data Poisoning Attacks)的鲁棒性。具体而言,针对MIA,应衡量攻击者推断某个样本是否参与训练的成功率,目标是使其接近随机猜测水平;针对数据投毒,则应评估在存在一定比例恶意客户端的情况下,模型性能下降的幅度,并设定可接受的性能损失阈值。

在数智商务场景下,版权归属与保护的完善是确保数字内容安全流通、促进商业创新的关键环节。尽管OpenAI等人工智能企业已经采用数字水印和区块链等技术为版权归属提供直接证据,但是仍然存在水印可能因内容篡改或格式转换失效,高并发交易场景下,公有链的吞吐量不足等问题。因此,数智商务场景下的版权保护应融合数字水印、内容指纹与区块链技术等技术优点,构建“技术嵌入—链上存证—智能管理”的体系。因此,数智商务场景下的版权保护应融合数字水印、内容指纹与区块链技术等技术优点,构建“技术嵌入—链上存证—智能管理”的体系。具体而言,数字水印的嵌入应关注其鲁棒性指标,即在内容经历常见格式转换、压缩或轻微编辑后,水印仍能被有效检测的概率,目标是确保在不影响感知质量的前提下,水印达到高水平的抗攻击能力;内容指纹技术应关注其唯一性识别率和抗碰撞性,即能够精确识别内容的独特性,并有效区分高度相似但非完全相同的内容,以支持精细化的版权追溯;区块链存证则应评估其交易吞吐量和确认延迟,选择能够满足高并发业务需求(如每秒处理数千笔甚至上万笔交易)且交易快速最终确认的分布式账本技术,以确保版权信息的及时记录和不可篡改性。

大模型遭受的攻击主要为提示注入和越狱攻击。提示注入是指通过构造特殊提示词绕过模型的安全对

齐机制,诱导模型执行非法操作;越狱攻击则是指利用模型的生成特性,突破预设的伦理或安全限制。这些攻击都是针对大模型特性进行的,Xie等[32]提供了一个包含各种类型的越狱提示和恶意指令的越狱数据集有效评估大模型抵御攻击的能力,但是在针对提示注入和越狱攻击等大模型攻击手段建立全流程威胁检测框架仍然存在不足。为构建可执行的防御体系,在威胁检测层面,应部署多阶段检测机制,包括输入端基于启发式规则、语义分析和少量样本学习的提示词过滤,以及输出端基于模型行为异常检测和有害内容分类器进行实时审查。关键指标包括检测率(Recall)和误报率(False Positive Rate),目标是显著提升对已知和未知攻击的覆盖度,同时将正常请求的误判率降至最低;在防御策略层面,可采用"系统模式自提醒"等方法,通过在系统层面注入安全指令,引导模型保持负责任的行为。其有效性可量化为攻击成功率(Attack Success Rate,ASR)的降低幅度,目标是使攻击成功率相对于无防御基线显著下降,同时需评估其对正常用户查询的响应质量和效率的影响,确保功能性不受损害。

AI算法决策的公平性是指算法决策出现决策偏差。尽管AI算法为商务场景带来了极大的便利,但是Celiktutan等[33]发现当算法的训练数据存在人类决策偏见时,算法也会出现相应偏差;Ren等[34]也发现真实世界的长尾数据也会导致训练数据和测试数据的不匹配问题。因此解决AI算法决策的公平性问题应聚焦AI鲁棒性与公平性,研究对抗样本防御,落地群体公平、机会均等等指标,解决长尾分布下的偏差问题。具体而言,在群体公平性方面,应选取并量化不同敏感属性群体之间的决策差异,例如比较其虚假阳性率(FPR)、虚假阴性率(FNR)或预测准确率等指标,使这些差异保持在可接受的最小范围内,从而实现"机会均等"或"预测平等"等公平性准则。在长尾分布问题上,应通过样本重加权、数据增强、少样本学习或元学习等方法,提升模型在稀有类别上的性能指标(如 F1 分数、精确率和召回率),从而缩小长尾类别与头部类别之间的性能差距,以确保算法决策对所有群体的覆盖和准确性,避免因数据稀疏导致的系统性偏差。

2.2 数智商务系统的智能协同与自主决策

数智商务系统的智能协同与自主决策聚焦的核心问题,是在人工智能与商务场景深度融合背景下,如何协调多智能体在商务生态中的合作博弈关系、实现人智协同决策的有效融合并构建适配智能体自主性的社会治理框架。例如,在技术快速进步的商务场景中,异构系统的数据孤岛、传统决策模式的滞后性等问题逐渐显现,导致需求预测、资源配置等环节的效率瓶颈。在此背景下,通过知识图谱、深度强化学习、联邦学习等技术路径,赋予系统动态感知环境、自主协调资源的能力,成为破解"信息孤岛化、决策滞后化、资源配置低效化"的关键。相关研究表明,人智协同中技能差异与信任机制影响显著,而智能体数量增加带来的合规性风险(如伦理与安全隐患),则需将透明性、责任性等伦理原则嵌入技术标准与法律制度,构建协同治理体系。

因此,围绕"多智能体博弈平衡—人机决策融合—智能体治理创新"逻辑链条研究系统智能协同与自主决策能力构建,是应对数字经济复杂性、推动产业数字化转型的重要路径,对提升商务系统效率、优化人机协作模式、防范技术应用风险具有理论与现实意义。

2.2.1 *多智能体决策的技术、方法与行为规律*

智能体作为能够感知环境并采取行动以实现特定目标的代理,其具备自主性、适应性和交互性等特征。其可以感知环境中的变化,并根据从环境学习到的知识和算法进行判断和决策,进而执行动作以影响环境或达到预定的目标。因此在众多关于智能体研究中,对其决策的技术、方法与行为规律的研究就尤为重要。

随着数据存储与计算能力的提升和以注意力机制为代表的大模型框架的提出[35],把人工智能作为研究对象的智能体决策研究逐渐成为主流。Lu等[36]发现移动设备和互联网的高度普及所带来的数字化用户行为数据,极大增强了人工智能决策的准确度。研究发现通过智能手机活动对用户金融风险进行画像在提升金融包容性方面的效果是使用在线社交媒体信息的1.3倍。Chakraborty等[37]开发了一种针对销售人员招聘的智能体决策模型,通过包含文本、语音和肢体语言等多模态信息的对话式视频面试录制资料,模型可以提取出客观且有理论支撑的销售绩效指标。该模型具有双向的动态交互能力与环境的实时适应能力,并且可以理解人类肢体语言,显著提高了对候选人的潜在销售能力的评估准确度。Yin等[38]考虑如何n将人工智能决策有效整合到常规诊断流程中,以人工智能决策的呈现时机为例进行研究。结果表明呈现时机对诊断的准确性具有显著影响,其事后建议表现最好,不提供人工智能建议的对照组表现最差。Chen等[39]发现智能体的训练数据本身可能存在偏差,因此将人工智能用于实际决策情境和工作流程之前,需要判断其是否表现出如基础率忽视、风险厌恶、确认偏误等行为决策偏差。通过对比发现,所使用的人工智能模型在近一半的实验中呈现出与人类相似的偏差行为,而且其在偏好驱动情境中的行为偏差更明显。

随着人工智能技术向复杂系统渗透,多智能体决策与博弈研究成为数智商务系统的核心议题。相较于单体智能体聚焦单一目标优化的局限性,多智能体系统需在动态环境中处理多个智能体间的交互关系—既包括合作场景下的任务协同与资源共享,也涵盖竞争场景下的策略对抗与利益博弈。这种复杂性使其广泛应用于策略博弈中的人机对抗、智能制造中的车间调度、智能交易中的高频策略互动等前沿领域。从理论演进看,该领域经历了从博弈论与符号逻辑的早期融合,到数据驱动的深度学习范式革新,再到认知科学与自然语言处理的跨学科突破,形成了多层次的研究体系。

早期多智能体研究融合博弈论与人工智能规划思想,以符号逻辑为工具构建理论模型。Rosenschein[40]在“理性互动”框架中探讨智能体间的交互策略,去除了分布式人工智能先前研究中的一些限制,通过形式化逻辑语言描述智能体的偏好、信念与决策规则,为博弈论与人工智能的交叉研究奠定基础。Zlotkin和Rosenschein[41]聚焦合作域中自主智能体间的谈判机制,针对分布式人工智能领域中智能体如何有效交互的问题,运用博弈论语言分析了离散域及智能体仅有部分信息情况下的谈判情形。区别于博弈论通常关注的连续域及全信息场景,该研究结合人工智能实际需求,提出一种新颖且稳定的谈判协议,用于解决智能体间离散任务的共享问题。

随着深度学习技术发展,多智能体决策进入以深度强化学习为核心的阶段。Jaderberg等[42]提出基于种群的强化学习方法,通过双层优化过程,让独立的强化学习智能体种群在随机生成的环境中,基于数千场并行匹配进行同步训练。每个智能体自主学习内部奖励信号,成功使智能体达到人类水平的游戏表现,展现出多智能体强化学习在复杂多人交互场景中的巨大潜力,为多智能体决策研究提供了新的方向。Ketter等[43]针对销售管理场景,提出一种结合经济环境特征与未来经济状况预测能力的实时智能体系统架构。该架构能够利用历史数据的统计分析,实时刻画观察到的市场状况,并预测不同时间尺度下市场状况的演变,以此做出战术(如产品定价)和战略(如成品库存水平)决策。展现出在零售行业的库存管理与定价决策中的应用潜力,凸显了多智能体系统在经济决策中的实际价值。

近年来,研究进一步融合自然语言处理与认知建模技术,推动智能体实现语言交互与意图理解。Xu等[44]针对基于LLMs的智能体在复杂决策任务中存在的固有偏差问题,提出一种将强化学习(Reinforcement Learning,RL)与LLMs相结合的新框架。通过LLM进行演绎推理生成多样化候选动作,再由RL策略优化决策,从而提升智能体的策略性和决策能力,实验证明该智能体在游戏中表现优于现有同类智能体。Hong等[45]为解决多智能体处理复杂任务时因幻觉问题导致的协作失败,提出MetaGPT框架,将人类有效工作流程以元编程方式融入基于LLMs的多智能体协作中,通过将标准操作流程编码为提示实现结构化协调,并要求模块化输出,赋予智能体领域专业知识以减少错误。在软件工程协作任务实验中,该框架相比现有多智能体系统能生成更连贯全面的解决方案。

2.2.2 人智协同决策与信任机制

随着人工智能技术在商业决策中的深度渗透,人智协同决策模式在数智化系统中越来越重要。其核心在于通过任务分工与互补机制,实现人类智能与机器算力的有机融合。自Licklider[46]首次提出了人机共生的概念,强调人类与计算机应通过紧密协作实现互补增效,而非简单的工具使用或者完全依赖于计算机。之后人机交互的研究经历了长期的发展,Zhang等[47]探讨了人机交互研究在商业应用和管理导向的独特性,为管理信息中人机交互的研究奠定了理论基础。在构建人智协同模式的过程中,仍有几个重要的问题需要解决:(1)在什么样的人智协同模式下生产力能得到提升;(2)AI的输出对于决策者来说都是黑盒,如何构建信任评估机制使AI做出的决策更加可信。

良好设计的人智协同模式能够促进生产力的提升。Fügener等[48]研究了在图像分类任务中,AI将任务委托给人类和人类将任务委托给AI两种任务委托的效率表现。研究发现,AI能精准识别自身的弱点,将自己不擅长的任务委托给人类,帮助提高总体效率。而人类无法准确判断何时要把任务委托给AI。因此设计人智协同的系统中,考虑AI对人类的委托更加重要。对于不同技能水平的工作人员,人工智能对其提高生产力的帮助不完全相同。Jia等[49]通过在电话营销公司的田野实验中发现对于高技能水平的员工来说,让AI生成销售线索能够显著提高销售转化率,而对于低技能水平的员工来说仅能小幅提升成功率。而在机器进行决策的过程中,人类的加入也能对机器的预测能力进行改善。如Lu和Zhang[50]在小额贷款行业中进行田野实验研究,考虑了不同信息复杂度与是否提供AI的特征解释对销售员业绩的影响。其研究发现在高信息复杂度,即提供了购物行为,线下轨迹等非传统数据,并且提供AI决策的特征解释时,人类介入能显著降低小额贷款公司的违约率。

AI建议对人类生产力是把双刃剑。人类与AI进行协作不能让人类的决策完全依赖于AI,否则可能影响人

类的创新性。Chen和Chan[51]指出非专家使用顾问模型能够创作出接近专家文本的水平。Fügener等[52]研究认为,AI建议能够保证个体的准确率得到提升,但可能导致人类的预测趋同于算法结果,降低人类群体决策的多样性。研究进一步提出可以增强AI输出结果的透明度,例如显示AI输出结果的置信度和局限性,能帮助人类判断何时应该依赖AI。

尽管人智协同的方式能够提升数智化系统的效率,但由于AI具有黑箱性,甚至可能输出不准确的结果,比如大语言模型存在幻觉,可能会输出虚假和错误的结果[53],因此人们对AI做出的决策仍需进行考量。那么在构建数智化过程中,形成人与AI之间的信任机制至关重要。为了解决上述问题,Afroogh等[54]系统性地梳理了构建可信AI的可操作框架,提出可以使用安全性,准确性,鲁棒性等技术性指标和伦理合规,法律合规,社会接受度等非技术性指标相结合的方式构建信任评估体系。而Balakrishnan等[55]的研究指出,人类在面对算法建议时,存在朴素建议加权(Naïve advice weighting, NAW)的现象,也就是人们会采用的固定权重对算法预测和自身价值进行加权,忽略私人信息价值的动态变化。那么人们在做出决策时可能会过度依赖算法或过度忽略算法的结果。考虑利用特征透明度对人类对AI存在的进行修正,则可以降低NAW的影响。

2.2.3 智能体社会治理机制的理论建构

随着人工智能技术的迅速发展,智能体社会治理机制的理论建构已逐渐成为跨学科研究领域的核心议题。尽管全球已初步达成了人工智能伦理原则的基本共识,但如何将伦理原则与法律制度、技术标准及治理实践有效协同,仍然面临诸多挑战。已有的研究围绕伦理原则的内在局限性、组织层面的治理创新机制、全球监管合作路径以及通用人工智能的风险控制问题展开了深入探讨,揭示出理论倡导与现实实践之间的复杂张力。

原则导向的治理框架在智能体社会治理研究中获得了广泛关注。当前,人工智能伦理原则在全球范围内表现出较高的趋同性,但其实际治理效力却受到多种因素的制约。Jobin等[56]通过研究全球的伦理指南发现,透明性、公正性、责任性、隐私保护与避免伤害五大原则趋于共识。然而,这些原则在具体解释逻辑、适用情境与落地实施策略等关键维度上仍存在明显差异。因此,仅依靠原则宣示难以保证治理成效,必须将原则制定与深度伦理分析及可操作的治理策略有机结合。Mittelstadt[57]指出,尽管人工智能伦理原则与医学伦理四原则高度相似,但人工智能领域缺乏统一的行业目标、职业规范和法律问责机制。这些高阶原则往往掩盖了深层的政治与规范分歧,难以转化为实践。

为解决原则治理的不足,许多学者提出应将伦理原则具体化为更为完善的社会法律治理框架。Theodorou和Dignu[58]认为,伦理原则必须嵌入具体的技术标准、法律制度和教育政策中,通过技术标准制定(如算法可解释性分级)、责任立法(如开发者产品连带责任)与伦理教育体系重构的三维联动,实现从原则宣示到行为规约的转化,同时明确人类在智能系统全周期中的责任边界。Zaidan和Ibrahim[59]从构建国际法与技术监测体系的角度,试图处理单一国家在技术能力不足和监管惯性导致“边境缺失困境”的问题,亟需通过国际法框架构建与技术权威网络重塑来破解集体行动困境。在生物安全领域,Bloomfield等[60]指出人工智能在生物技术领域的双重用途风险,这种特性使得传统自愿性伦理承诺效力锐减,导致可能被用于合成高致病性病原体。因此,需要政府实施强制性立法,将伦理原则转化为可操作的技术标准与法律条款,必须建立强制性生物模型风险评估框架,并将核酸合成供应链纳入全球技术治理体系,以避免可能引发全球疫情的重大风险。

组织层面的责任型人工智能治理机制创新成为近年研究热点。Berente等[61]智能体管理的三维分析框架,强调需在技术维度平衡自主性(Autonomy)与控制力,在认知维度破解黑箱化(Inscrutability)带来的责任真空,在演进维度管理持续学习(Continuous Learning)引发的系统不确定性,并通过信息系统管理框架实现协调与控制。Hadley等[62]从理论上回顾了算法审查委员会的实践经验。跨行业案例研究表明,算法审查委员会的有效性取决于领导层政治承诺、现有制度流程的适应性改造以及伦理成本内部化机制的协同作用,单纯技术审查难以应对商业利益最大化的制度惯性。Gans[63]通过博弈模型发现,即使存在以社会福祉为导向的资助方,也难以在市场竞争中遏制营利性企业开发的非受限通用人工智能产品。这一结论挑战了传统“善意治理”假设,提示通用人工智能治理需超越企业自律框架,探索跨国强制性干预路径。

2.3 数智商务变革的新模式与新业态

随着大模型与人工智能技术的飞速发展和深度融合,数智商务正以赋能所带来的产品服务创新与智能运营模式的重塑。然而,商业活动的这些变革与创新并非孤立存在,它们迫切需要与之相适应的、有效的组织结构与市场机制提供坚实的保障。

数智商务变革的新模式与新业态主要分为三个层面:(1)针对AI赋能的商务流程创新,主要研究企业商务流程的两大核心阶段:一方面,在新产品开发阶段,关注

多模态交互驱动的客户体验价值创造等关键问题；另一方面，在运营管理阶段，深入研究效率—韧性双目标协同优化等核心流程优化。(2)针对AI驱动的组织变革，主要研究现代组织的形态与运作模式的转型，尤其关注AI驱动的员工认知与行为调适、人智协作模式与人智分工优化、员工绩效极化、组织AI应用战略等方面。(3)针对人智融合新市场的机制设计，主要关注“人+AI”多主体共创型市场的演化：重新审视人智融合下的新型竞争机制、人智协同的激励机制设计、人智共创的价值分配机制，以及人—机—平台三元制度设计等核心问题。

2.3.1　AI赋能的商务流程创新

AI技术正在重构传统商务流程的运行逻辑。从供应链到客户服务，AI不仅通过自动化提升流程效率，更通过数据驱动的决策模式创造新的流程价值增长点。已有研究主要从供应链与运营、企业决策、创意与研发、客户交互等维度的流程探讨了AI如何重塑商务流程的底层逻辑。图3展示了AI驱动下数智商务新模式与新业态的三个层次。

(1)供应链与运营流程

在供应链方面，Cen等[64]研究了AI技术的采纳在供应链中的扩散，发现客户对AI的采纳会显著带动上游供应商对AI的采纳，AI采纳的扩散显著提升了供应商的质量管理水平并降低其运营成。进一步发现，高AI暴露员工的流动以及地理距离较近等因素会强化AI技术在供应链中的扩散效应；表明AI技术采纳在供应链中的扩散过程是通过知识和经验的传播而促进AI的应用(即“学习渠道”)，而非单纯为了迎合客户需求(即“迎合渠道”)。在运营流程方面，Choi等[65]探讨了“先发货后购物”的订阅模式下企业如何借助AI实现利润最大化。研究发现，企业在该模式中的策略会随着AI对客户需求的预测能力的变化而呈现非单调变化。当AI预测能力较低时，企业倾向于通过“高订阅费+低价格”的策略来提取事前剩余；当AI预测能力处于中等水平时，企业策略转变为“低订阅费+高价格”以提取事后剩余；而AI预测能力极高时，由于消费者展示购物风险(即消费者利用订阅服务发现产品后转向其他低价渠道购买)增加，企业又会重回高订阅费策略。这些研究表明，AI在供应链与运营流程中的应用，需结合行业特性与实际业务场景，充分发挥其优化资源配置和决策支持的作用。已有研究在运营方面还探讨了AI在产品推荐方面的应用。Longoni和Cian[66]提出了“机器口碑效应”(Word-of-Machine Effect)，发现消费者因AI与人类在功利性和享乐性评估能力上的认知差异而对两类推荐产生偏好分化；消费者购物时在功利性目标下更偏好AI推荐，这是因为消费者认为AI更理性和高效；在享乐性目标下消费者更偏好人类推荐，因为消费者认为人类更懂情感体验。上述研究为企业理解AI在供应链中的扩散，以及根据AI的能力和运营场景优化AI使用策略提供了理论依据。



图3 数智商务变革的新模式与新业态的三个层次

Fig.3 Three Levels of New Models and Formats of Digital Commerce Transformation

(2)企业决策流程

在企业决策层面，已有研究主要探讨了企业在AI算法选择方面的决策和AI辅助企业进行重要的业务决策。在AI算法选择方面，企业在众多的AI算法选择时会面临算法偏差和算法方差方面的平衡。算法越复杂(比如使用更多消费者特征变量)，越能够精准拟合数据(低偏差)，但过度拟合也可能导致预测结果不稳定(高方差)；相反，简单算法虽可以降低方差而提升预测稳定性，但也可能因为预测精度不够而导致预测偏差。Lyer和Ke[67]研究市场竞争如何影响企业的算法选择，发现垄断企业倾向选择复杂算法，追求精准预测从而识别高价值客户。这是因为，即使算法预测结果波动大(即高方差)，垄断企业也不必过于担心失去客户；相反，市场竞争会促使企业选择简单算法模型，因为复杂算法可能让双方都瞄准同一高价值群体，导致激烈竞争，从而降低企业利润，而简单算法通过减少变量使企业更可能均匀覆盖不同群体，降低目标重叠，软化市场竞争。在AI算法辅助企业决策方面，AI的采纳使得传统商务流程忽视的“数据尾迹”正成为AI创新的富矿。例如，Li等[68]关于普惠金融的研究发现，传统的信用传统金融机构(如银行)在评估贷款申请时，严重依赖信用历史、收入证明等“强信号”数据，但弱势群体(如低收入人群、农村居民、个体经营者)往往缺乏这些数据，导致贷款通过率低，陷入“无信用记录→无法贷款→更难建立信用记录”的恶性循环；基于AI信用评分模型可以利用传统模型忽略的“弱信号”数据(如社交媒体行为、手机使用记录、社保公积金缴纳情况等)，并通过机器学习算法挖掘复杂的数据关联，更精准地评估信用风险，可能打破传统金融的歧视性壁垒。

(3)创意与研发流程

在创意与研发流程方面,Noy和Zhang[69]发现使用生成式AI工具可显著提升中级专业写作任务(如撰写新闻稿、撰写简短报告、制定分析计划、撰写内容正式或敏感的电子邮件)的生产力;这些任务的平均完成时间在使用生成式AI工具后减少40%,而输出质量提升18%。Chen和Chan[51]基于广告文案的创作研究发现,顾问模式可以显著提升非专家的文案质量,使其接近专家水平,而代笔模式会降低专家文案创作的质量。这是由于AI工具代笔起草的文案可能会对专家的创作产生锚定效应,从而使得创作内容出现同质化倾向。Lou和Wu[70]基于AI在生物制药企业的应用提出了AI创新能力(AI Innovative Capacity, AIIC)的概念,通过对生物制药企业的专利和相关数据的深入分析,发现AIIC能够显著提升临床前药物—靶点对的发现效率。这种提升效果对于作用机制已知、化学新颖性中等的药物效果最为显著,而对全新机制或增量药物的效果则相对有限。

(4)客户交互流程

AI已经被广泛应用于客户交互的流程,其中AI客服的拟人化设计是现有文献中探讨的重要议题之一。例如,Schanke等[71]的研究表明,赋予AI虚拟形象与情感语调(如共情式回应)等拟人化特征,可使客户留存率提升超过20%;但拟人化程度的提高也可能使得客户对价格敏感程度提高,这是因为用户将拟人化的机器视为"人",容易激发其与人类商量和谈判的心理。这种客户转化率与价格敏感度之间的矛盾说明企业在应用AI与客户进行交互时应根据企业的目标导向而对AI进行适当调整。研究还发现机器和人工在不同类型的交互场景中各有优势。例如,Wang等[72]在语音呼叫中心场景中,对高频查询类问题(如账单查询)采用纯机器应答,而对投诉类问题切换至"人工坐席+AI辅助"的混合模式,可以使平均处理时间缩短31%的同时,客户满意度提升8%。Garvey等[73]研究了消费者对AI与人类传递的正负向信息存在差异反应,结果表明,负面信息(如低于预期的报价)由AI传递时,消费者接受度更高,这是因为消费者认为AI不存在自私意图;而正面信息(如高于预期的报价)由人类传递时,消费者满意度更高,这是因为消费者感知人类有更强的善意。研究还发现,拟人化AI会强化意图感知,使负面信息接受度下降、正面信息满意度上升。上述研究表明,AI工具与客服互动过程中需要在拟人化与效率之间进行权衡,同时AI和人类在不同的服务类型中各有优势。

2.3.2 AI驱动的组织变革

随着AI技术的快速发展及其在组织中的广泛应用,AI对组织的影响已超越工具层面,正在重塑组织的员工行为、能力体系、权力结构和组织战略。这种变革的本质是组织认知框架的重构,即当AI从执行工具逐步升级为决策主体,传统的科层制、专业化分工、绩效评估体系均面临重大挑战。已有文献主要从员工认知与行为调适、员工与组织绩效极化、人智协作与分工优化以及组织的AI应用战略等方面探讨了AI如何颠覆组织运行的底层逻辑。

(1)员工认知与行为调适

AI算法被越来越多地引入到组织管理的过程中,员工也在经历人工管理到算法管理的时代,其在认知和行为层面都需要重新调适,由此出现了员工对AI算法的信任、抵制和厌恶等问题。在AI算法信任方面,Li和Bitterly[74]发现,当组织用AI直接评估员工绩效时,即使算法能够做到绝对公平,员工信任度仍比人类评估低29%,这种信任度缺失的原因在于员工对AI算法的善意感知的缺失(即认为AI缺乏对员工主观关怀意图)。对AI和算法的信任危机可能引发员工对AI的隐性抵制。例如,SimanTov-Nachlieli等[75]在针对销售团队的研究中发现,高绩效员工对AI推荐客户的抵触率高达67%,原因是这些员工更加自信"比算法更懂客户",并且担心由于AI的采用而使得自身的优势地位丧失。类似的,Liu等[76]对网约车司机的大规模研究也发现,司机对算法的接受度受个体情境经验(如历史接单量)和从众行为影响;当算法推荐与司机在特定时段和地点的经验冲突,或同行选择与算法相悖时,司机更可能拒绝推荐。Turel和Kalhan[77]探讨人类对AI的隐性偏见与算法厌恶的动态关系,发现人们普遍存在对AI的隐性不信任偏见,且这种偏见在参考点为"人类专家"时更强;这种对算法的偏见可以显著预测个体的算法厌恶;随着人们对AI性能信息的获取,偏见对决策的影响逐渐减弱。这些研究表明,组织在实施AI管理或引入AI算法时,不仅要关注AI的技术效能,更要重视员工的情感体验和情景需求,构建基于信任的人智协同管理模式,消除员工的AI算法的抵制和偏见。

(2)员工与组织的绩效极化

已有文献在多个领域的研究都发现,AI技术的使用可能会在个体和组织层面引发绩效的分化。在个体层面,Hui等[78]追踪自由职业平台的设计师对AI的使用发现,AI工具使高技能设计师收入增长32%,但初级设计师接单量下降41%。类似的,Krakowski等[79]也发现,掌握AI协作技巧(如提示词工程)的员工,其问题解决速度比仅依赖传统技能的同事快2.3倍。这些研究表明AI的使用能让高技能水平的员工受益更多,从而导致员工之

间绩效的差异增加。在组织层面,Otis等[80]探讨生了成式AI(GPT-4驱动的WhatsApp助手)对企业绩效的影响,发现高绩效企业家使用AI后绩效提升15%,而低绩效企业家的绩效则下降8%;原因在于高绩效者能有效筛选AI建议并实施针对性策略,而低绩效者倾向采用通用建议,因未适配自身需求导致负向效果。与上述研究发现不同,Gaessler和Piezunka[81]通过分析1970—2000年超2万名棋手的50万场比赛数据,发现国际象棋计算机(AI)作为训练伙伴可显著提升棋手水平,尤其使低水平棋手和训练机会少的棋手受益,缩小技能差距。Noy和Zhang[69]在探讨使用生成式AI工具对写作任务绩效的影响时也发现,使用AI工具可以显著提高写作任务的生产力,其中能力水平较低的员工更能够从使用生成式AI工具中获益,显示出了生成式AI工具在中级专业写作任务中的使用可以减少因员工能力差异而导致的生产力不平等。上述研究表明,并不是所有的个体都能够从AI使用中获得均等的收益,AI使用对个体绩效的影响不仅取决于是否使用,还可能取决于如何使用以及人智协作的模式。

(3)人智协作模式与人智分工优化

人与AI协作过程中的不同人智协作模式及其对绩效的影响也是现有研究探讨的重点议题之一。例如,Boussioux等[82]基于可持续循环经济商业创意众包挑战赛,对比了人类群体与两种人机协作模式(即独立搜索和差异化搜索)生成的解决方案;研究结果发现,人类群体产生方案的平均新颖程度更高,但人机协作方案在战略可行性、财务与环境价值及整体质量上更优,其中差异化搜索(即人类引导AI迭代生成不同方案)能在不牺牲上述价值的前提下提升新颖性。Lebovitz等[83]在AI赋能医疗诊断场景中提出了整合型增强(Engaged Augmentation)和非整合型增强(Unengaged Augmentation)两种人智协作模式;基于美国某医院放射科的田野研究发现,在肺癌、乳腺癌和骨龄诊断三个场景中,仅肺癌诊断科室通过与AI的有效交互实现整合型增强,将AI结果与专业判断结合,降低不确定性;而乳腺癌和骨龄诊断科室因时间压力、工具复杂性及结果不可解释性等原因呈现非整合型增强,要么忽视AI结果,要么盲目接受。已有文献还探讨了人智的不同竞争优势及其互补关系。例如,Krakowski等[79]基于国际象棋比赛场景研究AI对竞争优势来源的影响,发现AI的引入可以引发替代和互补动态效应,即传统人类棋力(如Elo等级分)的竞争优势被AI削弱(替代效应),但人类与AI协作的新能力(如选择和调优引擎)成为新优势来源(互补效应),并且这类新能力与传统能力无显著关系,表明AI驱动竞争优势从人类单一能力转向人机协作能力。

(4)组织的AI应用战略

除个体层面之外,已有研究还探讨在组织层面探讨了AI的应用战略。例如,Raisch和Krakowski[84]提出组织在AI应用上面临自动化—增强化(Automation-Augmentation)悖论;如果组织应用AI主要目标在于追求流程自动化(如用AI替代客服),AI应用在短期内会显著降低成本,但也会削弱组织学习能力(员工失去客户洞察机会);如果组织对AI应用战略侧重于增强模式(如AI辅助人类决策),AI应用可以保留组织的学习潜力,但需持续投入培训成本。Lebovitz等[83]探讨AI应用于医疗诊断领域的研究也解释了类似的矛盾,当AI仅提供辅助建议时,医生诊断准确率提升19%;但当AI完全接管初诊时,医生的病理分析能力在6个月内下降31%。解决上述悖论需要组织动态平衡自动化和增强化。Anthony等[85]主张将AI不仅仅视为辅助工作的工具或连接团队的媒介,而应将其视为写作的参与者,这就需要研究要研究AI与人类、组织、社会的动态关系网络,因为AI的影响不仅来自技术本身,还来自其所处的社会系统。在AI应用与组织适配方面,针对AI技术所具有的通用性(不是某个企业所独有)、透明性(数据和算法逻辑容易被窃取或模仿)、短视性(只专注于特定任务、缺乏对全局的理解)等特点,Kemp[86]提出情境化AI战略框架,主要内涵包括:扎根(Grounding),引导AI聚焦企业特有的数据和业务场景,形成差异化能力,而非简单的通用化学习;划界(Bounding),通过制度、技术手段限制上述AI差异化能力的扩散,防止竞争对手模仿;重塑(Recasting)调整AI算法和组织流程,使其适应企业战略和跨部门协作需求,让AI适配战略全局。

2.3.3 人智融合新市场的机制设计

随着大模型技术,特别是生成式AI与智能体技术的成熟与广泛应用,传统平台经济正加速向人类用户与AI智能体共同参与、协同创造价值的新型市场形态演化。这种人智融合新市场的出现,不仅重塑了商业逻辑与价值链,也对现有市场运行规则、资源配置和治理框架带来了深刻挑战。传统机制设计理论难以完全适用于包含具有不同认知、决策和学习能力的AI智能体的新场景。因此,为这一新兴市场设计有效、公平、能引导良性互动并促进价值共创的机制,已成为数智商务领域的关键科学问题。本节将围绕新型竞争机制、人智协同激励、共创价值分配及人—机—平台三元制度等核心议题,系统梳理研究进展,剖析挑战,展望未来。

(1)人智融合下的新型竞争机制

人智融合新市场的核心特征是人类与AI智能体作

为异质性参与者的竞争与合作。AI的引入根本性地改变了竞争优势的来源。Krakowski等[79]的研究表明,AI同时引发替代(AI替代人类传统能力)和互补(人机互动产生新兴“人机能力”)效应,后者成为新的竞争优势来源,且可能与传统能力不相关。这意味着市场竞争已演变为“人—AI混合体”之间的协同效能竞争。在此背景下,理解人智主体间的动态互动逻辑至关重要。Rai等[87]提出的“人—AI混合体”谱系为理解不同层次的人智互动提供了框架。数字平台通过任务模块化、数据生成和AI学习模态为这些混合体运作提供了基础,但也可能带来平台利用AI强化主导地位或AI偏见影响公平竞争等新问题。Li等[88]通过斯塔克尔伯格博弈模型展示了多主体异质市场中的分层战略互动。聚焦人智主体间的动态互动逻辑以揭示协同与排他机制的形成与演化,能够为多元主体共存提供理论支撑。未来需进一步研究不同类型人—AI混合体的竞争行为及促进公平竞争的市场规则。

(2)人智协同的激励机制设计

促进人智融合市场从共存走向深度协同,关键在于设计有效的激励机制。面对AI的自主性、学习能力和人类创造性的复杂性,传统激励理论亟需拓展。AI既可被激励,也可作为激励工具。Han等[89]通过演化博弈模型指出,奖惩机制均能影响AI行为,但效果与风险各异,提示激励设计需精细权衡。Inkpen等[90]强调,人—AI团队绩效提升的关键在于用户专业知识与AI算法的互补性调优,激励机制需识别并奖励这种“人机互补性”。如前所述,Jia等[49]的研究揭示了AI辅助对人类员工创造力的提升作用具有显著的“技能偏向性”,即高技能员工能更有效地利用AI辅助来处理重复性任务,从而将更多精力投入到更高层次的创新性问题解决中,实现更大的创造力提升。这启示研究者在设计人智协同的激励机制时,不仅要关注AI任务完成质量与人类创意参与度,更需要充分考虑人类参与者的技能水平差异。现有研究着重构建以AI任务完成质量与人类创意参与度为核心的新型协同激励框架,研究激励的动态调整、公平收益共享,并兼顾效率、可解释性与策略透明度。未来研究应探索针对不同技能水平人类用户的差异化激励策略并促进其创造力的发展。此外,还需深入研究AI在激励机制设计与动态调整过程中的智能化应用,以实现更精准、高效的协同激励。

(3)人智共创的价值分配机制

在人智共创模式下,价值链涉及人、AI及平台等多方主体,其贡献量化与价值公平分配成为核心难题。传统价值分配理论难以直接适用。Krakowskti等[79]指出,AI的引入重塑了价值来源,竞争优势更多源自新兴的“人机能力”,价值度量需从个体贡献转向协同效能。如何设计既能准确反映各方贡献—AI的“智能贡献”和人类的“增强价值”,又能被广泛接受,并利用新技术保障分配过程可信的机制,是关键科学问题。Chowdhury等[91]的研究进一步强调了在AI与员工协作的环境中,知识共享、员工的AI技能、对AI的信任以及角色清晰度是影响协作效果并最终提升业务绩效的关键因素。这表明,在评估和分配人智共创价值时,不能仅仅关注最终产出,还必须考虑这些促进协作与价值实现的过程性因素和个体能力。生成式AI驱动下,商业模式和价值主张均需要重塑。平台作为价值网络的组织者,其角色至关重要。为此,未来应深入研究人—机—平台共创价值链中各参与方贡献的度量方法与公平分配原则。核心探索方向是结合区块链、智能合约等技术实现可信履约与数据溯源,设计透明高效的分润机制。Rai等[87]关于“数据监护”和基于区块链的自决主权数据监护模式的讨论,为此提供了理论借鉴。未来需开发新的价值度量模型和分配算法,并在不同共创场景中检验其适用性。

(4)人—机—平台三元制度设计

在人智融合新市场中,平台角色从传统中介转变为复杂人智协同生态的制度构建者与协调者,形成人—机—平台三元互动体系。这要求超越双边市场理论,构建能有效治理三方关系、平衡利益并确保系统稳定运行的新制度。下一代数字平台需解决人—AI混合体带来的偏见、数据监护和市场势力失衡等治理问题。例如,制度需确保市场“厚度”—足够的人与AI参与者,通过规则减少“拥堵”,通过信任机制和透明度保障“安全”,设计“稳定”规则防范机会主义,并考虑伦理价值避免“厌恶”交易。Han等[89]的研究表明监管和激励机制的设计对引导AI发展至关重要。Lee和Cui[92]针对零工平台争议解决机制的研究为这种三元制度设计提供了具体的例证。他们比较了平台中心化裁决与用户投票的去中心化裁决机制,发现去中心化机制能更有效地消除平台偏见,实现更公平的争议解决。这凸显了在人—机—平台三元体系中,将部分治理权下放给用户社群,通过精心设计的机制来协调各方利益的重要性。在此背景下,Chen等[93]进一步基于人—环境匹配理论,发现平台治理的正式化能直接促进零工的工作重塑;然而,平台工作不安全感氛围会削弱这种积极关系。这表明,在设计人—机—平台三元制度时,不仅要考虑争议解决机制的去中心化,更要关注治理规则的明确化与正式化,以及营造安全的工作环境,从而平衡平台、零工和客户间的利益,帮助零工更好地通过工作重塑来应对工作困境。

未来应更多研究如何设计支持多主体博弈与高效资源配置的三元制度框架,引入“可控自由度”概念,实现平台在激励约束、人机治理与责任归属间的精妙平衡[94],增强市场机制的长期稳态运行能力。

3　未来5~10年大模型时代数智商务的理论和方法发展目标及资助重点

3.1　发展目标

前沿技术的快速迭代与应用场景的交叉渗透,为数智商务发展提供了持续的动力、多样性和探索空间,也带来了知识的爆炸、复杂性和研究挑战。面对全球数智商务领域的研究竞争,大力发展面向数智商务的跨学科研究,建立面向数智商务的自主知识体系,是提升我国数字经济全球竞争力的重要举措。在技术层面,加速多元智能、人类增强、智能体等核心技术的突破,抢占世界科技前沿的制高点;在管理层面,形成领域导向的数智商务应用模式,构建完整的技术与经济闭环,从而在技术与规则的国际竞争中保持领先地位。在未来五到十年中,需要进一步凝练该研究领域的内涵、发展目标、核心科学问题、关键技术问题和典型应用场景,夯实学科理论基础、创新方法、培养具有前沿数智商务理论的人才,以应对当前大模型时代对数智商务人才的需求与技术需求。

3.2　资助重点

本次“双清论坛”的与会专家经过深入研讨,凝练了大模型时代数智商务的理论和方法的若干重大关键科学问题,并建议未来5~10年应着重围绕以下领域开展原创性研究。重点研究领域清单如表1所示。

(1)面向数智商务的关键技术与方法

背靠人工智能与大模型技术的快速发展,数智商务已成为推动数字经济高质量发展的重要引擎。商务场景中产生的多模态、高维度数据是优化商业决策的关键要素,而数据隐私与模型安全则是保障可持续发展的基础屏障。若能有效融合垂直领域知识,突破轻量化建模技术,构建安全可靠的计算框架,将充分释放商务数据的要素价值,实现智能技术与商业实践的深度融合。

因此,建议资助重点为:1)资源受限商务场景下人工智能算法研究:重点研究大模型的高效微调与推理优化技术,包括低秩分解、参数冻结等方法;开发模型压缩算法,如剪枝、量化等技术;探索轻量级架构设计和自动搜索方法;研究知识迁移与适配增强技术,提升小模型性能。2)面向数智商务的大模型垂域化关键技术研究:构建多模态数据治理与质量增强框架;研发分布式异构算力协同调度机制;设计通用知识与领域特征协同建模方法;完善商务智能体全生命周期管理技术。3)复杂商务决策场景下的大模型强化学习:开展多源数据驱动的环境仿真建模研究;开发高维状态空间下的分层强化学习算法;设计多目标奖励函数;研究虚实环境安全迁移方法。4)领域增强的商务智能技术与方法:研究结构化知识嵌入技术,融合行业规则与专家经验;开发领域理论引导的模型架构优化方法;探索小样本场景下的领域适配预训练技术。5)商务智能领域的安全与风险管理

表1　未来5~10年重点研究领域清单

Table 1　List of Key Research Areas for the Next 5~10 Years

关键科学问题	技术与方法缺口	潜在资助方向
面向数智商务的关键技术与方法	大模型垂域化与AI智能体技术	资源受限商务场景下人工智能算法研究
		面向数智商务的大模型垂域化关键技术研究
	AI增强的下一代商务智能技术	复杂商务决策场景下的大模型强化学习
		领域增强的商务智能技术与方法
	商务智能安全与风险技术	商务智能领域的安全与风险管理关键技术
数智商务系统的智能协同与自主决策	智能体决策理论与方法	智能体决策理论与跨领域应用创新
		多智能体协同博弈与复杂系统建模
		人智协同决策机制与信任体系构建
	智能体社会行为与治理	智能体社会治理框架与伦理合规体系
		商务智能领域的安全与风险管理关键技术
数智商务变革的新模式与新业态	AI赋能的商务流程创新	多主体供应链协同决策的智能化管理机制
		人智协同决策中的权力分配与组织效能提升
		基于“弱信号”挖掘的智能决策范式重构
	AI驱动的组织变革	算法管理下员工行为调适与组织干预策略
		AI驱动组织战略的动态适配与生态治理
	人智融合新市场机制	人智融合新市场的核心机制设计与演化规律研究

关键技术:突破隐私计算技术,优化联邦学习框架;研发数字版权保护方案;开发模型安全防御机制,包括对抗攻击检测等内容;建立公平性保障体系,实现偏差缓解与合规评估。

(2)数智商务系统的智能协同与自主决策

随着人工智能与商务场景的深度融合,数智商务系统正经历从集中式决策向智能协同决策的革命性转变。在此过程中,智能体的自主决策能力提升、多主体协同机制构建、人机信任体系完善及跨层面治理框架设计等问题,成为制约数智商务可持续发展的核心瓶颈。为推动理论创新与技术突破,需聚焦智能体决策逻辑、多主体博弈建模、人智协同机制、社会治理体系及生态效能优化等关键领域,系统性布局研究方向。

因此,建议重点资助方向如下:1)智能体决策理论与跨领域应用创新:以有限理性模型为基础,结合大模型框架探索智能体决策逻辑,推动其在营销、金融、医疗等领域的融合应用,如基于多模态数据的销售人员能力评估模型、信用风险评估中的AI决策优化,同时关注智能体决策偏差的识别与修正。2)多智能体协同博弈与复杂系统建模:突破动态竞争与协作场景下的博弈理论,融合博弈论与深度强化学习,研究自动驾驶、智能制造等场景中智能体的交互机制,同时探索自然语言处理与认知建模在多智能体协作中的应用。3)人智协同决策机制与信任体系构建:分析不同技能水平人员与AI的协同模式,突破信任机制技术瓶颈,结合准确性指标与伦理合规性构建评估体系,并明确人智协同中的责任划分与风险防控。4)智能体社会治理框架与伦理合规体系:推动透明性、公正性等伦理原则向技术标准与法律制度转化,探索跨国治理协作与责任型AI治理模式,应对智能体自主性与不可解释性带来的监管挑战。5)动态商务生态中的决策效能优化:平衡多智能体生态中个体效益与系统整体效能,开发兼顾竞争与协作的激励算法,优化“人机—机机”混合协同网络,借鉴销售管理中的实时决策系统架构,结合边缘计算等技术提升分布式决策效率,应对市场化开发与公共利益的潜在冲突。

(3)数智商务变革的新模式与新业态

由AI技术突破和广泛应用驱动的数智商务正在重塑商业逻辑的底层架构,学术界亟需跨越技术应用的表象,深入探讨数智商务引发的组织认知框架的重构、价值创造范式的颠覆和市场竞争规则的迭代,亟需探索数智商务变革的新模式与新业态,为企业在技术跃迁中实现价值跃升提供理论支撑。

围绕“流程重构—组织变革—市场创新”的递进逻辑,建议资助重点为:1)多主体供应链协同决策的智能化管理机制:针对供应链中企业、供应商、客户等多方利益冲突问题,研究AI如何构建动态协同决策模型。重点探索AI在需求预测、库存优化和风险预警中的协同作用,开发基于博弈论和契约设计的智能协调算法,解决信息不对称下的利益分配难题。2)人智协同决策中的权力分配与组织效能提升:围绕AI在企业战略决策中的角色演化,研究人类管理者与AI系统的权责边界划分。结合组织行为学理论,开发“人类主导—AI验证”或“AI提议—人类否决”等混合决策模式。3)基于“弱信号”挖掘的智能决策范式重构:研究AI对决策相关“弱信号”的提取与识别,构建决策增强系统,研究弱信号在风险预警、需求洞察中的决策增益,突破制度障碍,设计数据治理框架与决策文化培育机制,为企业实现“全息决策”提供方法论支持。4)算法管理下员工行为调适与组织干预策略:针对AI算法引发的员工抵触行为,研究员工心理抵抗的微观机制,识别算法透明度、反馈机制设计对员工信任的影响路径,开发“算法共治”管理模式,提出兼顾效率与人性化的组织制度设计方案,为企业缓解人机矛盾提供实践指南。5)AI驱动组织战略的动态适配与生态治理:从动态能力理论出发,研究企业如何通过AI技术吸收实现能力升级。构建“AI技术—组织学习—竞争优势”理论框架,探索AI与组织战略的互动机制,提炼AI能力培育的关键节点,研究AI生态中的竞合规则与治理框架,提出防范技术垄断、促进中小企业创新的政策建议,为国家产业升级、数字经济治理体系建设提供参考。6)人智融合新市场的核心机制设计与演化规律研究:整合动态博弈、激励设计、价值分配理论与技术,深入研究人智异质主体在新兴市场中的互动、合作与价值共创机制,针对人—机—平台三元互动特性,探索数据治理、算法伦理、竞争政策、权益保护以及“可控自由度”等新型制度安排,确保市场长期稳态运行与可持续发展。

4 结 语

大模型时代数智商务中蕴含着很多新的管理学问题,为管理学科的发展提出了新的要求与新的方向。数智商务正成为科技与产业深度融合的关键前沿。我国超大规模的市场、完备的产业体系、庞大的数据资源和丰富的应用场景也为数智商务研究提供了广阔的探索机会。运用数字技术和智能算法,对企业的业务流程、运营模式、决策机制等进行全面的升级和再造,驱动新质生产力的发展,需要同时具备技术的突破和管理的创新。根据相关研究热点与趋势,本文讨论了面向商务场景人工智能模型与技术的部署、增强与风控;人与智能

体、多智能体的全链路协同与自主决策;以及数智商务在流程、组织、市场三个层面的理论与应用创新。本文凝练出了该领域未来五到十年的重大关键科学问题,探讨了前沿研究方向和科学基金资助战略。今后需要进一步深入研究大模型时代数智商务理论和方法中的核心科学问题,推动范式跃迁,构建具有自主性的理论体系,推动构建智能、高效、协同的数智商务系统新格局,助力我国数智经济高质量发展,提升我国数字经济全球竞争力。

参 考 文 献

[1] 姜富伟,刘雨旻,孟令超. 大语言模型、文本情绪与金融市场. 管理世界,2024,40 (08):42—64.
Jiang FW,Liu YM,Meng LC. Large language model and textual sentiment analysis in Chinese stock markets. Journal of Management World,2024,40 (08):42—64. (in Chinese)

[2] 李浩,邢志远,李琳利,等. 基于多智能体的工业数字孪生系统云边端架构与关键技术. 计算机集成制造系统,2024,30 (11):3755—3770.
Li H,Xing ZY,Li LL,et al. Cloud-edge-device system architecture and key technologies of industrial digital twin system based on multi-agent. Computer Integrated Manufacturing Systems,2024,30 (11):3755—3770. (in Chinese)

[3] 于嘉汀,冷嘉承,徐雪晴,等. 连锁零售企业供应链智能补货系统研究. 运筹与管理,2024,1—9.
Yu JT,Leng JC,Xu XQ,et al. The research on the intelligent replenishment system of the supply chain in chain retail enterprises. Operations Research and Management Science,2024,1—9. (in Chinese)

[4] 杜梦瑶,李清明,张淼,等. 面向隐私保护的用户评论基准数据集构建与大模型推理能力评估. 计算机学报,2025,48(07):1529—1550.
Du MY,Li QM,Zhang M,et al. Constructing benchmark datasets for privacy - protected user comments and evaluating the reasoning capabilities of large models. Chinese Journal of Computers,2025,48 (07):1529—1550. (in Chinese)

[5] Porter ME,Heppelmann JE. How smart,connected products are transforming competition. Harvard Business Review. 2014,92(11):64—88.

[6] Constantinides P,Henfridsson O,Parker GG. Introduction—platforms and infrastructures in the digital age. Information Systems Research,2018,29(2):381—400.

[7] 陈剑,刘运辉. 数智化使能运营管理变革:从供应链到供应链生态系统. 管理世界,2021,37 (11):227—240+14.
Chen J,Liu YH. Operations management innovation enabled by digitalization and intellectualization:from supply chain to supply chain ecosystem. Journal of Management World,2021,37 (11):227—240+14. (in Chinese)

[8] Toorajipour R,Sohrabpour V,Nazarpour A,et al. Artificial intelligence in supply chain management:A systematic literature review. Journal of Business Research,2021,122:502—517.

[9] Lyu ZY,Lin P,Guo DQ,et al. Towards zero-warehousing smart manufacturing from zero-inventory just-In-time production. Robotics and Computer-Integrated Manufacturing,2020,64:101932.

[10] 宋华. 人工智能数智供应链的理论探索与展望. 中国流通经济,2024,38 (01):44—54.
Song H. Ai - based intelligent supply chain:Theoretical exploration and prospects. China Business and Market,2024,38 (01):44—54. (in Chinese)

[11] Hu EJ,Shen Y,Wallis P,et al. Lora:Low-rank adaptation of large language models.// International Conference on Learning Representations. 2022.

[12] Saha R,Srivastava V,Pilanci M. Matrix compression *via* randomized low rank and low precision factorization.//Advances in Neural Information Processing Systems. 2023.

[13] Sun M,Liu Z,Bair A,et al. A simple and effective pruning approach for large language models.//International Conference on Learning Representations. 2023.

[14] Tan M,Le Q. Efficientnet:Rethinking model scaling for convolutional neural networks.//International Conference on Machine Learning. 2019. PMLR.

[15] Hao Z,Guo JY,Han K,et al. One-for-all:Bridge the gap between heterogeneous architectures in knowledge distillation.// Advances in Neural Information Processing Systems. 2023.

[16] He J,Zhou CT,Ma XZ,et al. Towards a unified view of parameter-efficient transfer learning.//International Conference on Learning Representations. 2021.

[17] Li WB,Zhou H,Yu JQ,et al. Coupled mamba:Enhanced multi-modal fusion with coupled state space model. arXiv preprint,2024,arXiv:2405.18014.

[18] Mirhoseini A,Goldie A,Yazgan M,et al. A graph placement methodology for fast chipdesign. Nature,2021,594(7862):207—212.

[19] Zhang JQ,Zhao Y,Saleh M,et al. Pegasus:Pre-training with extracted gap-sentences for abstractive summarization.//International Conference on Machine Learning. 2020. PMLR.

[20] Zheng JH,Qiu SJ,Shi CM,et al. Towards lifelong learning of large language models:A survey. ACM Computing Surveys,2025,57(8):1—35.

[21] Huang WP,Jiang W,Luo XG,et al. It's showtime:Live-streaming E-commerce and optimal promotion insertion policy. Production and Operations Management,2024:10591478231224975.

[22] Yan XL,Chen YF,Yu H,et al. Newsvendor problems with product unbundling:An approach combining robust optimization with deep reinforcement learning. Production and Operations Management,2025,34(11):3629—3646.

[23] Chen J,Xu YF,Yu PW,et al. A reinforcement learning approach for hotel revenue management with evidence from field experiments. Journal of Operations Management,2023,69(7):1176—1201.

[24] Dubois Y,Li CX,Taori R,et al. Alpacafarm:A simulation framework for methods that learn from human feedback. Advances in Neural Information Processing Systems,2023. 36:30039—30069.

[25] Li GR. Construct to Associate:Cooperative context learning for domain adaptive point cloud segmentation.// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.

[26] Yu ZY,Shen L,Ding L,et al. Sheared backpropagation for fine-tuning foundation models.// Proceedings of the IEEE/CVF Conference on

Computer Vision and Pattern Recognition. 2024.

[27] Liu JH,Zhang CC,Guo JY,et al. Ddk:Distilling domain knowledge for efficient large language models. Advances in Neural Information Processing Systems,2024. 37:98297—98319.

[28] Wu S,Wang Y,Yao Q,Why in-context learning models are good few-shot learners?// International Conference on Learning Representations. 2025.

[29] Price WN II,Cohen IG. Privacy in the age of medical big data. Nature Medicine,2019,25(1):37—43.

[30] Dwork C,Roth A. The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science,2014,9(3/4):211—407.

[31] Zhou JX,Chen SY,Wu YL,et al. PPML-Omics:A privacy-preserving federated machine learning method protects patients' privacy in omic data. Science Advances,2024,10(5):eadh8601.

[32] Xie YQ,Yi JW,Shao JW,et al. Defending ChatGPT against jailbreak attack *via* self-reminders. Nature Machine Intelligence,2023,5(12):1486—1496.

[33] Celiktutan B,Cadario R,Morewedge CK. People see more of their biases in algorithms. Proceedings of the National Academy of Sciences of the United States of America,2024,121(16):e2317602121.

[34] Ren JW,Yu CJ,Ma X,et al. Balanced meta-softmax for long-tailed visual recognition.// Advances in Neural Information Processing Systems. 2020.

[35] Vaswani A,Shazeer N,Parmar N,Attention is all you need.// Advances in Neural Information Processing Systems. 2017.

[36] Lu T,Zhang YJ,Li BB. Profit *vs*. equality? The case of financial risk assessment and a new perspective on alternative data. MIS Quarterly,2023,47(4):1517—1556.

[37] Chakraborty I,Chiong K,Dover H,et al. Can AI and AI-hybrids detect persuasion skills? Salesforce hiring with conversational video interviews. Marketing Science,2025,44(1):30—53.

[38] Yin JM,Ngiam KY,Tan SS,et al. Designing AI-based work processes:How the timing of AI advice affects diagnostic decision making. Management Science,2025:mnsc.2022.01454.

[39] Chen Y,Kirshner SN,Ovchinnikov A,et al. A manager and an AI walk into a bar:Does ChatGPT make biased decisions like we do? Manufacturing & Service Operations Management,2025,27(2):354—368.

[40] Rosenschein JS,Rational interaction:Cooperation among intelligent agents. United States:Stanford University,1986. .

[41] Zlotkin G,Rosenschein JS,Negotiation and task sharing among autonomous agents in cooperative domains.// International Joint Conference on Artificial Intelligence. 1989.

[42] Jaderberg M,Czarnecki WM,Dunning I,et al. Human-level performance in 3D multiplayer games with population-based reinforcement learning. Science,2019,364(6443):859—865.

[43] Ketter W,Collins J,Gini M,et al. Real-time tactical and strategic sales management for intelligent agents guided by economic regimes. Information Systems Research,2012,23(4):1263—1283.

[44] Xu Z,Yu C,Fang F,Language agents with reinforcement learning for strategic play in the werewolf game. arXiv preprint, 2023,arXiv:2310.18940.

[45] Hong S,Zhuge M,Chen J,et al. MetaGPT:Meta programming for a multi-agent collaborative framework.// International Conference on Learning Representations. 2023.

[46] Licklider JCR,Man-computer symbiosis. IRE Transactions on Human Factors in Electronics,1960. 1:4—11.

[47] Zhang P,Nah FF,Benbasat I. Special section:Human-computer interaction research in management information systems. Journal of Management Information Systems,2005,22(3):9—14.

[48] Fügener A,Grahl J,Gupta A,et al. Cognitive challenges in human–artificial intelligence collaboration:Investigating the path toward productive delegation. Information Systems Research,2022,33(2):678—696.

[49] Jia N,Luo XM,Fang Z,et al. When and how artificial intelligence augments employee creativity. Academy of Management Journal,2024,67(1):5—32.

[50] Lu T,Zhang YJ. 1 + 1 > 2? Information,humans,and machines. Information Systems Research,2025,36(1):394—418.

[51] Chen ZN,Chan J. Large language model in creative work:The role of collaboration modality and user expertise. Management Science,2024,70(12):9101—9117.

[52] Fügener A,Grahl J,Gupta A,et al. Will humans-in-the-loop become borgs? Merits and pitfalls of working with AI. MIS Quarterly,2021,45(3):1527—1556.

[53] Huang L,Yu WJ,Ma WT,et al. A survey on hallucination in large language models:Principles,taxonomy,challenges,and open questions. ACM Transactions on Information Systems,2025,43(2):1—55.

[54] Afroogh S,Akbari A,Malone E,et al. Trust in AI:progress,challenges,and future directions. Humanities and Social Sciences Communications,2024,11(1):1568.

[55] Balakrishnan M,Ferreira KJ,Tong J. Human-algorithm collaboration with private information:Naïve advice-weighting behavior and mitigation. Management Science,2025:mnsc.2022.03850.

[56] Jobin A,Ienca M,Vayena E. The global landscape of AI ethics guidelines. Nature Machine Intelligence,2019,1(9):389—399.

[57] Mittelstadt B. Principles alone cannot guarantee ethical AI. Nature Machine Intelligence,2019,1(11):501—507.

[58] Theodorou A,Dignum V. Towards ethical and socio-legal governance in AI. Nature Machine Intelligence,2020,2(1):10—12.

[59] Zaidan E,Ibrahim IA. AI governance in a complex and rapidly changing regulatory landscape:A global perspective. Humanities and Social Sciences Communications,2024,11(1):1121.

[60] Bloomfield D,Pannu J,Zhu AW,et al. AI and biosecurity:The need for governance. Science,2024,385(6711):831—833.

[61] Berente N,Gu B,Recker J,et al. Managing artificial intelligence. MIS Quarterly,2021,45(3):1433—1450.

[62] Hadley E,Blatecky A,Comfort M. Investigating algorithm review boards for organizational responsible artificial intelligence governance. AI and Ethics,2025,5(3):2485—2495.

[63] Gans JS. Can socially minded governance control the artificial general intelligence beast? Management Science,2025,71(10):8188—8199.

[64] Cen L,Han Y,Qiu J et al. Artificial intelligence along the supply chain. Market Intelligence,2024.

[65] Choi WJ,Liu QH,Shin J. Predictive analytics and ship-then-shop subscription. Management Science,2024,70(2):1012—1028.

[66] Longoni C,Cian LC. Artificial intelligence in utilitarian *vs*. hedonic

contexts:The “word-of-machine” effect. Journal of Marketing,2022,86(1):91—108.

[67] Iyer G,Ke TT. Competitive model selection in algorithmic targeting. Marketing Science,2024,43(6):1226—1241.

[68] Li CX,Wang HC,Jiang ST,et al. The effect of AI-enabled credit scoring on financial inclusion:Evidence from an underserved population of over one million. MIS Quarterly,2024,48(4):1803—1834.

[69] Noy S,Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science,2023,381(6654):187—192.

[70] Lou BW,Wu L. AI on drugs:Can artificial intelligence accelerate drug development? Evidence from a large-scale examination of bio-pharma firms. MIS Quarterly,2021,45(3):1451—1482.

[71] Schanke S,Burtch G,Ray G. Estimating the impact of “humanizing” customer service chatbots. Information Systems Research,2021,32(3):736—751.

[72] Wang WG,Gao GG,Agarwal R. Friend or foe? Teaming between artificial intelligence and workers with variation in experience. Management Science,2023:mnsc.2021.00588.

[73] Garvey AM,Kim T,Duhachek A. Bad news? Send an AI. Good news? Send a human. Journal of Marketing,2023,87(1):10—25.

[74] Li MY,Bitterly TB. How perceived lack of benevolence harms trust of artificial intelligence management. Journal of Applied Psychology,2024,109(11):1794—1816.

[75] SimanTov-Nachlieli I. More to lose:The adverse effect of high performance ranking on employees’ preimplementation attitudes toward the integration of powerful AI aids. Organization Science,2025,36(1):1—20.

[76] Liu M,Tang XC,Xia SY,et al. Algorithm aversion:Evidence from ridesharing drivers. Management Science,2023:mnsc.2022.02475.

[77] Turel O,Kalhan S. Prejudiced against the machine? Implicit associations and the transience of algorithm aversion. MIS Quarterly,2023,47(4):1369—1394.

[78] Hui X,Reshef O,Zhou LF. The short-term effects of generative artificial intelligence on employment:Evidence from an online labor market. Organization Science,2024,35(6):1977—1989.

[79] Krakowski S,Luger J,Raisch S. Artificial intelligence and the changing sources of competitive advantage. Strategic Management Journal,2023,44(6):1425—1452.

[80] Otis NG,Rowan C,Solene D et al. The uneven impact of generative AI on entrepreneurial pserformance. Working Paper,2024.

[81] Gaessler F,Piezunka H. Training with AI:Evidence from chess computers. Strategic Management Journal,2023,44(11):2724—2750.

[82] Boussioux L,Lane JN,Zhang MM,et al. The crowdless future? Generative AI and creative problem-solving. Organization Science,2024,35(5):1589—1607.

[83] Lebovitz S,Lifshitz-Assaf H,Levina N. To engage or not to engage with AI for critical judgments:How professionals deal with opacity when using AI for medical diagnosis. Organization Science,2022,33(1):126—148.

[84] Raisch S,Krakowski S. Artificial intelligence and management:the automation–augmentation paradox. Academy of Management Review,2021,46(1):192—210.

[85] Anthony C,Bechky BA,Fayard AL. “Collaborating” with AI:Taking a system view to explore the future of work. Organization Science,2023,34(5):1672—1694.

[86] Kemp A. Competitive advantage through artificial intelligence:Toward a theory of situated AI. Academy of Management Review,2024,49(3):618—635.

[87] Rai A,Constantinides P,Sarker S,Next-Generation digital platforms:Toward Human-AI Hybrids. MIS Quarterly,2019. 43(1):iii-ix.

[88] Li JM,Ai Q,Yin SR,et al. An aggregator-oriented hierarchical market mechanism for multi-type ancillary service provision based on the two-loop Stackelberg game. Applied Energy,2022,323:119644.

[89] Han TA,Pereira LM,Lenaerts T,et al. Mediating artificial intelligence developments through negative and positive incentives. PLoS One,2021,16(1):e0244592.

[90] Inkpen K,Chappidi S,Mallari K,et al. Advancing human-AI complementarity:The impact of user expertise and algorithmic tuning on joint decision making. ACM Transactions on Computer-Human Interaction,2023,30(5):1—29.

[91] Chowdhury S,Budhwar P,Dey PK,et al. AI-employee collaboration and business performance:Integrating knowledge-based view,socio-technical systems and organisational socialisation framework. Journal of Business Research,2022,144:31—49.

[92] Lee WK,Cui Y. Should gig platforms decentralize dispute resolution? Manufacturing & Service Operations Management,2024,26(2):519—536.

[93] Chen JY,Luo JL,Hu WA,et al. Fit into work! From formalizing governance of gig platform ecosystems to helping gig workers craft their platform work. Decision Support Systems,2023,174:114016.

[94] Le Guillou M,Prévot L,Berberian B. Bringing together ergonomic concepts and cognitive mechanisms for human—AI agents cooperation. International Journal of Human–Computer Interaction,2023,39(9):1827—1840.

Digital Intelligence Business in the Era of Large Language Models: Theoretical Advances, Key Scientific Issues, and Methodological Approaches

Xin Xu[1] Qiang Ye[2] Hansheng Wang[3] Zhiguang Ren[4] Guangxin Jiang[5] Chaoyue Gao[2]
Hao Xia[5] Haitian Xie[3] Danqi Hu[3] Jialu Chang[1] Zhijie Lin[1*]

1. *School of Economics and Management, Tsinghua University, Beijing 100084, China*
2. *School of Management, University of Science and Technology of China, Hefei 230026, China*
3. *Guanghua School of Management, Peking University, Beijing 100871, China*
4. *Department of Management Sciences, National Natural Science Foundation of China, Beijing 100085, China*
5. *School of Management, Harbin Institute of Technology, Harbin 150001, China*

Abstract In response to national priorities and the frontiers of academic research, the development of digital and intelligent commerce in the era of large models depends not only on breakthroughs in key technologies such as artificial intelligence and their applications in the industry but also on the reengineering of traditional business processes and innovations in management theories. Deeply integrating data intelligence with business scenarios, overcoming core technological bottlenecks, fostering new industries and models, and achieving a new paradigm of intelligent, efficient, and collaborative digital and intelligent commerce represent crucial future directions. Based on the 406th "Shuang Qing Forum", focusing on the theories and methods of digital intelligence business in the era of large language models, this study delves into three aspects: key technologies and methods, intelligent collaboration and autonomous decision-making, and new models and forms. It further proposes a research agenda addressing advancements in AI methods and large model technologies, multi-agent decision-making and cross-agent collaboration, and human-AI integration in process-organization-market reengineering, along with funding recommendations prioritizing interdisciplinary research in management science and computational science for digital and intelligent commerce.

Keywords large language models; digital intelligence business; multi-agent; human-AI collaboration; governance system; organizational change

徐 心 清华大学经济管理学院副院长、史带讲席教授、人工智能与管理(AIM)研究中心主任。从事信息系统学科研究,主要研究领域包括信息技术(IT)商业价值与IT治理、数字化转型、商务计算与分析等。担任中国系统工程学会信息系统工程专业委员会(CNAIS)理事长、*Information Systems Research*(*ISR*)与*MIS Quarterly*(*MISQ*)等期刊Senior Editor。

林志杰 清华大学经济管理学院长聘副教授,中国信息经济学会青工委主任、管理科学与工程学会信息系统与数字化创新分会副主任委员。研究领域包括信息系统经济学、数字化平台管理等。担任*MIS Quarterly*(*MISQ*)Associate Editor及*Journal of the Association for Information Systems*(*JAIS*)Senior Editor。

(责任编辑 王 磊 张 强)

* Corresponding Author, Email: linzhj@sem.tsinghua.edu.cn