

DOI: 10.16538/j.cnki.fem.20240712.101

组织管理中的人工智能决策: 述评与展望

张亚莉, 李辽辽, 丁振斌

(西北工业大学 管理学院, 陕西 西安 710129)

摘要: 本文聚焦于组织管理领域的AI决策问题,对2018—2024年发表的176篇国内外文献进行了系统回顾,剖析组织决策中AI与人类的复杂互动关系。文章从决策要素、决策过程和决策结果三个维度揭示了AI决策与人类决策的本质差异,并提炼出两种AI与人类决策的交互模式:替代决策和合作决策。替代决策的相关研究主要关注不同决策主体之间的差异及其引发的潜在影响;合作决策的相关研究则更加关注合作机制设计、AI对合作决策的影响与合作效果的优化。在此基础上,按照“决策模式选择—决策过程—决策结果”的逻辑构建了一个AI赋能下的综合决策框架,充分考虑了替代决策与合作决策的重叠性和互补性,并强调了人与AI协同的重要性。最后,从AI决策的前因后果、决策机制设计以及中国特殊情境的影响三个方面对未来研究进行了展望,为推动AI从低阶感知智能向高阶决策智能的跃迁提供参考。

关键词: 人工智能;替代决策;合作决策;人机协同

中图分类号: F270 **文献标识码:** A **文章编号:** 1001-4950(2024)10-0018-21

一、引言

人工智能(artificial intelligence, AI)在组织决策中的应用一直是AI领域最重要的研究话题之一(Cao, 2022)。专家预测,第四次工业革命将由数字化、信息和通信技术、机器学习、机器人和AI推动,并将更多的决策权从人类转移到机器上(Syam和Sharma, 2018)。相较于传统决策支持工具,AI具有模拟甚至超越人类认知能力的潜质,能够执行特定决策任务,并通过对大规模数据的深入分析来改善和加速决策过程(Metcalf等, 2019)。当前,基于AI技术的决策(简称“AI决策”)已经在人力资源(Langer等, 2023)、医疗保健(Gaube等, 2021)、创新管理(Pietronudo等, 2022)、战略管理(Wu等, 2023)等诸多领域得到了广泛应用。

AI的参与改变了以人为中心的传统决策模式(王锋和刘玮, 2023),其在更广泛的组织决策任务中不断取代和补充人类行动的现象引起了学术界的高度关注(Dwivedi等, 2021; Krakowski等, 2023)。随着超级计算机和大数据技术的迅猛发展,AI执行任务的认知能力不断提高,有机会承担更复杂的决策任务(Dwivedi等, 2021; Sharma等, 2024)。部分学者认为凭借愈

收稿日期: 2023-08-21

作者简介: 张亚莉(1972—),女,西北工业大学管理学院教授,博士生导师;

李辽辽(1998—),女,西北工业大学管理学院博士研究生(通讯作者, liliaoliao@mail.nwpu.edu.cn);

丁振斌(1996—),男,西北工业大学管理学院博士研究生。

发复杂的数据来源和算法模型, AI能够成为替代人类决策的有效工具(Araujo等, 2020; Langer和Landers, 2021)。其他学者则对人类与AI的合作潜力表现出了相当大的兴趣(Choudhary等, 2023; 张志学等, 2024), 主张将AI视为增强人类能力的工具, 而非人类的替代品(Jarrahi, 2018), 倡导从更积极角度探究AI与人类如何更好地协同决策(Tschandl等, 2020)。

尽管AI在组织决策领域的应用前景广阔, 但也伴随着诸多争议, 如何在复杂决策情境下有效发挥人—AI协同决策的潜力尚未明晰(Shrestha等, 2019)。AI决策因其结果随时间动态变化、使用过程不透明以及运算逻辑难以理解的特性, 常被使用者视为“黑匣子”(Anthony等, 2023)。这导致了决策过程的不透明性和决策结果解释性的缺失(孔祥维等, 2021), 严重降低了AI决策的可信度(Chen等, 2024; Jussupow等, 2021)。引发了人们对AI决策的复杂情感, 既有对其高效和精确性的欣赏, 也有对其不可预测性和潜在风险的担忧, 形成了算法欣赏与算法厌恶两种截然不同的态度(魏昕等, 2021; 张志学等, 2024)。对于重要的组织决策, AI决策的合法性问题也被质疑, 尤其在涉及伦理道德和算法责任等敏感领域时(Waldman和Martin, 2022)。此外, AI决策可能存在的偏见和歧视问题也引起了学者们的广泛关注(肖红军, 2022)。

针对AI在组织决策领域的作用及影响, 国内外学者进行了初步探索(Jarrahi, 2018; 裴嘉良等, 2021), 但在组织管理的背景下, 尚缺乏对AI决策文献的系统性梳理, 限制了学者对AI在组织决策中的具体作用及其与人类决策者互动关系的全面理解。有必要重新审视AI介入的组织决策范式, 建立一个综合性的“人—AI决策”研究框架来帮助研究者更好地进行问题定位, 厘清已有研究结论的冲突, 从而更好地提供实践指导。

本文通过系统梳理组织管理中AI决策的相关文献, 深入剖析了AI作为决策主体参与组织决策的具体作用及其潜在影响, 并指出了潜在的研究方向。文章结构安排如下: 首先, 梳理出176篇国内外相关文献用于研究分析; 然后, 从多维度揭示AI决策和人类决策的根本区别, 明确AI与人类决策的两种交互模式: 替代决策和合作决策; 接着, 总结提炼两种交互模式下各自领域的研究重点; 在此基础上, 按照“决策模式选择—决策过程—决策结果”的分析逻辑, 构建AI赋能下的综合决策框架。最后, 从多角度指出了未来可能的研究方向。

二、文献收集与计量分析

(一) 文献收集

首先, 在数据收集之前界定了AI决策的概念边界。AI决策是指通过识别、解释、推断并从数据中学习, 从而提出决策或行动方案, 以实现预定的组织和社会目标的自动化过程(Mahmud等, 2023)。与传统的决策支持系统不同, AI决策系统可以自主学习、适应和行动(Pathirannehelage等, 2024)。在没有AI介入的决策系统中, 人类负责识别并定义与决策相关的各类因素; 而在有AI的决策系统中, 算法根据从数据中识别出的模式和规律, 自主决定决策应当考虑的因素(Mahmud等, 2022)。组织管理中的AI决策, 特指通过引入AI技术, 辅助或替代人类决策者, 对组织运营和管理等关键决策问题进行分析、判断和提出优化方案。

然后, 明确文献的收集范围, 包括在组织管理情境下详细描述了AI决策的过程和方法、明确使用了AI技术进行决策, 以及对比人类决策和AI决策等相关的文献。需要说明的是, 本研究不涉及具体算法设计和方法的技术优化, 不区分专家系统、机器学习、遗传算法、神经网络等不同AI算法得到结果的差异。

最后, 界定检索词和检索范围。关键词的初始设定根据作者讨论形成, 通过初始关键词检索筛选的文献分析, 形成第二次检索的关键词, 按照这种方式进行关键词的迭代更新, 直到不再产生新的关键词。迭代形成的关键词包括主要人工智能决策、AI决策和算法决策。“人工智能

决策”和“AI决策”直接关联到本研究的核心研究问题。“算法决策”被纳入检索词,是因为一些研究虽采用“算法决策”的命名方式,但实质上更多地关注基于AI的算法决策。如景怡(2023)研究中的算法决策是指以大数据为基础,以人工智能技术为核心,通过算法处理做出的决策。

为揭示最新研究进展,本文将检索时间限定在2018年1月1日到2024年5月30日之间。其中,中文文献来源于中国知网,检索关键词:主题=“人工智能决策”OR“AI决策”OR“算法决策”(文献类型为期刊,来源类别为CSSCI);外文文献来源于Web of Science核心数据库和FT50期刊,主要检索关键词包括“AI decision-making”“Artificial Intelligence decision-making”“Algorithmic decision-making”“Human machine decision-making”“Organizational management”等(文献类型为Article,来源类别为SCI和SSCI)。

详细检索流程见图1。检索过程共划分为三个阶段。第一阶段在Web of Science核心数据库、FT50期刊和中国知网数据库中直接搜索关键词,并将检索结果根据研究领域、标题进行筛选,去掉非组织场景以及和主题没有直接关联的文献,初步筛选得到886篇、80篇和145篇相关文献。第二阶段通过阅读关键词和摘要,删除只讨论人工智能、算法,但与决策关联较弱的文献。然后,合并重复文献,进一步去除关联性较弱的文献,最终形成文献214篇。第三阶段通过阅读全文,删除只在摘要中提及决策,但核心内容与AI决策相关度不高的文献。最终确定43篇中文文献和133篇外文文献,共176篇文献用于研究分析。

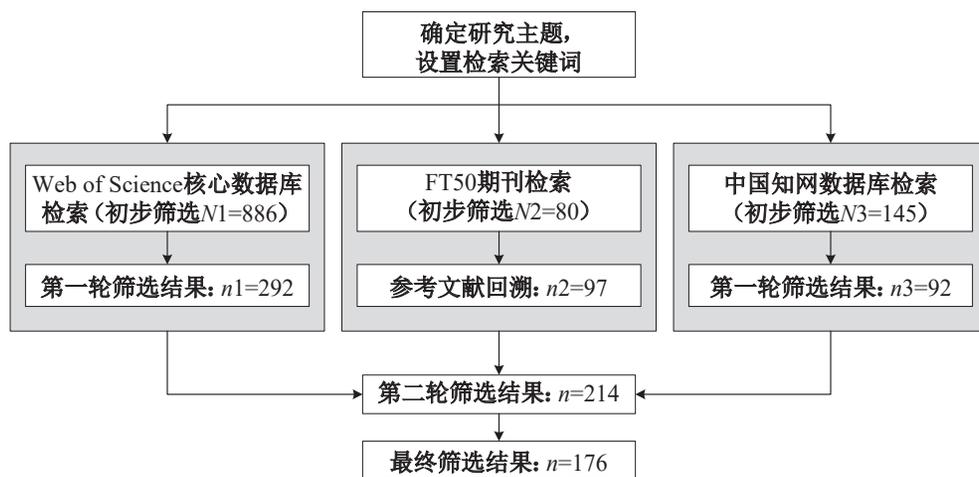


图1 文献检索流程

(二)文献计量分析

1. 整体发文趋势

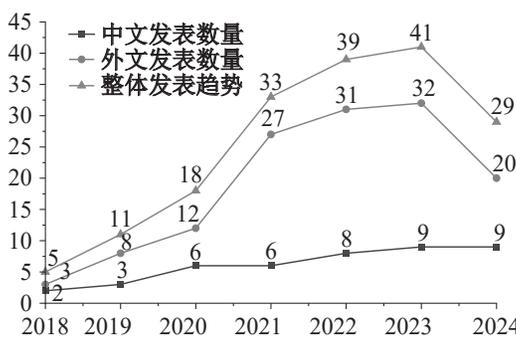
外文文献主要集中在知名的科技和管理类期刊,如*Computers in Human Behavior*、*Management Science*、*Information Systems Research*、*Technological Forecasting and Social Change*、*Journal of Management Information Systems*、*International Journal of Information Management*等。中文文献主要发表在《心理科学进展》《经济管理》《心理学报》《管理科学》《中国人力资源开发》《管理世界》《外国经济与管理》等期刊。图2列出了国内外发文数量和趋势。可以发现,组织管理领域的AI决策研究逐渐获得更多研究者的关注,近几年的文献数量呈现不断上升的趋势。2021年是关键的转折点,文献数量开始急剧增加,之后持续上升。这表明研究者们对组织领域中AI决策问题的高度关注。

2. 研究作者和研究领域分布

对国内外文献发表的研究作者和研究领域分析发现:(1)从作者分布来看,美国的发文量高达50篇,远超其他国家和地区,在AI决策领域的研究处于引领地位。之后是德国、中国和加拿大,发文数量在10~20篇之间,也具有较强的研究实力和活跃度。此外,荷兰、英国、澳大利亚、法国等国家也表现出较大的研究兴趣,发文数量在3~10篇之间。中文文献的作者所在机构主要是武汉大学和清华大学,发文数量均为6篇。(2)从作者影响力来看,外文最具

代表性的学者是Langer Markus、Dwivedi Yogesh和Duan Yanqing,发文量均为3篇;中文文献最具代表性的作者是清华大学的许丽颖和武汉大学的喻丰,发文量分别达到4篇和3篇。(3)从研究领域来看,相关研究主要聚焦在商业经济学领域,心理学的研究数量紧随其后。信息科学与图书馆科学、运筹学与管理科学、计算机科学、工程学以及社会科学等领域也呈现出显著的研究兴趣,反映出组织管理中的AI决策研究具有明显的学科交叉特性。

相关研究涵盖了多种决策场景,具体见表1。



注:文献检索时间跨度为2018年1月1日至2024年5月30日。

图2 组织管理中AI决策研究发文量年度分布及趋势

表1 组织管理中AI决策的应用场景

| 应用场景 | 决策类型 |
|------|---|
| 人力资源 | 招聘决策(Choung等,2023;Langer等,2021)、求职(Wesche和Sonderegger,2021)、人力资源决策(Langer等,2023;Yan等,2024;裴嘉良等,2021)、员工绩效反馈(Yam等,2022)等 |
| 供应链 | 采购决策(Nissen和Sengupta,2006)、客户关系管理(Gaczek等,2023)、零售商决策(Li和Li,2022;Sharma等,2024;Tschandl等,2020)等 |
| 医疗 | 医疗保健决策(Longoni等,2019;Park等,2022)、临床决策(Komorowski等,2018;Liang和Xue,2022)、医疗诊断决策(Jussupow等,2021)等 |
| 营销 | AI营销代理(Liu-Thompkins等,2022)、服务机器人(Castelo等,2023;Luo等,2019)、投资推荐(Chua等,2023)、专业服务(Spring等,2022)、AI推荐(Logg等,2019;杨炳成等,2023)、新闻报道检测(Santoro和Monin,2023)等 |
| 财务 | 会计决策(Estep等,2024)、贷款决策(Costello等,2020;Sachan等,2020)等 |
| 生产 | 任务分类(Fügener等,2022)、订单包装(Sun等,2022)等、产品分类(Kawaguchi,2021)、任务分配(Bai等,2022)等 |
| 其他 | 道德决策(Gill,2020;Zhang等,2022)、不确定性决策(Dietvorst和Bharti,2020;Liu,2021)、公共决策(Grimmelikhuijsen,2023)、自动化领导决策(Höddinghaus等,2021)、战略决策(Keding和Meissner,2021)等 |

3. 领域研究热点

通过剖析关键词的共现关系,构建出该领域研究的热点图谱。使用VOS Viewer对外文文献的高频关键词进行共现聚类分析,统计出现频率在5次及以上的关键词,得到图3;对中文文献的高频关键词进行共现聚类分析,统计出现频率在3次及以上的关键词,得到图4。

从图3可以看出,国外组织管理中的AI决策研究中,高中心度的关键词主要包括:artificial intelligence、trust、management、impact、automation和performance等。通过对文献中高频关键词的共现聚类分析,可将组织中的AI决策研究归纳为四个研究热点:(1)探讨AI决策在伦理道

德、公平正义以及偏见等方面的本质性问题；(2)关注人工智能决策应用的核心挑战；(3)关注AI决策应用的影响；(4)关注人们对AI决策的态度。

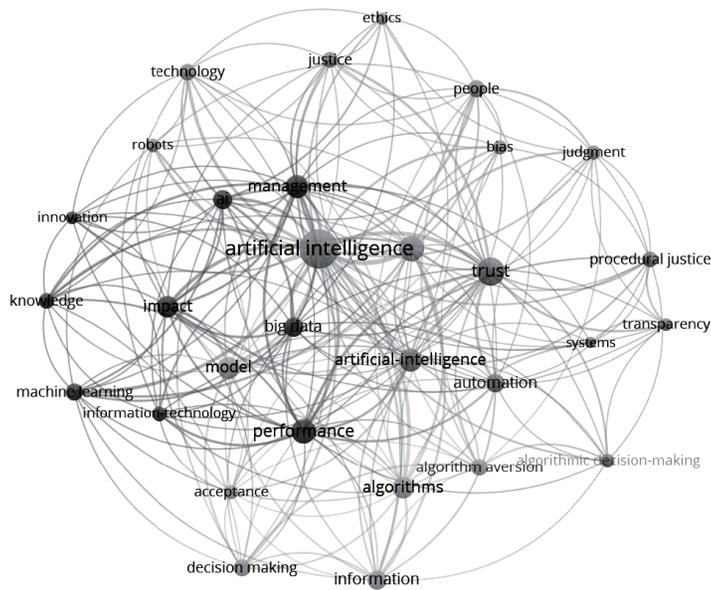


图3 国外文献高频关键词共现知识图谱

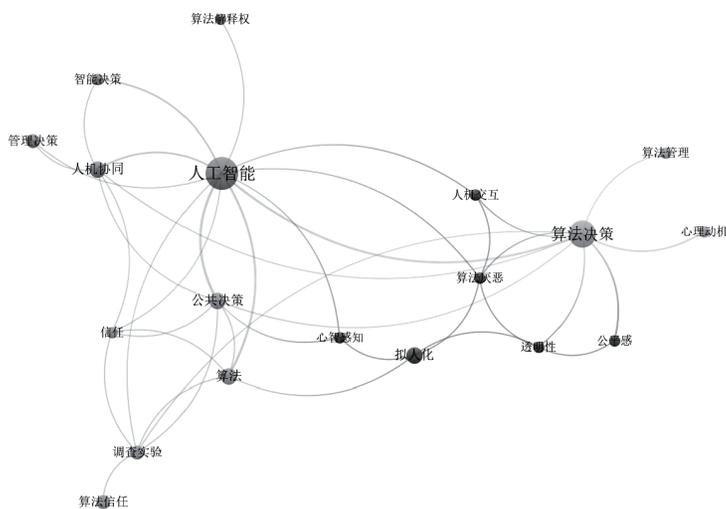


图4 国内文献高频关键词共现知识图谱

从图4可以看出,国内组织管理中的AI决策研究中,高中心度的关键词包括:人工智能、算法决策、公共决策、人机协同、算法信任和算法厌恶等。以人工智能为中心形成的聚类显示了学者们对人机协同下AI决策的重视。公共决策相关的高频关键词是信任,与人工智能和算法决策关联的高频关键词是人机交互、算法厌恶、透明度、公平感、拟人化和心智感知,而与算法决策联系紧密的还有算法管理和心理动机。

根据Shrestha等(2019)的研究,人工智能时代的组织决策结构可划分为四种类型:对AI完全授权(Full human to AI delegation)、AI与人类按顺序决策(Hybrid 1: AI to human sequential

decision-making)、人类与AI按顺序决策(Hybrid 2: Human to AI sequential decision-making)、聚合人类与AI(Aggregated human-AI decision-making)。

综合国内外文献高频关键词的聚类分析,本研究从“人”与“AI”两个决策主体之间的关系入手,将组织管理中的AI决策归纳为两个核心主题:(1)AI代替人类决策(替代决策):研究主要围绕AI代替人类决策的影响因素(如高频关键词:automation、trust、justice、transparency、拟人化、公平感等)和决策后果(如高频关键词:performance、innovation、ethics、bias等)展开。这些研究对应于“Full human to AI delegation”决策类型。(2)AI与人类合作决策(合作决策):研究主要集中在AI与人类合作决策的过程及其影响(如高频关键词:人机交互、人机协同、management、算法管理、心理动机)。这些研究对应于“Hybrid 1: AI to human sequential decision-making”“Hybrid 2: Human to AI sequential decision-making”和“Aggregated human-AI decision-making”三种决策类型。

三、AI决策与人类决策

(一)AI决策与人类决策的区别

AI决策和人类决策既具有密切的联系又形成了不可忽视的差异。一方面,AI起源于对人类智能行为的模仿,二者在基于理性思考的决策中具有相似性,均致力于达成“最优”决策。而且AI决策系统具备自主学习和演进的能力,这是它们与人类决策过程相似且优于传统决策系统的重要特征。另一方面,人和被视为“机器”的AI在实体类型以及心理距离等多方面存在不同,人们很难将“换位思考”“推己及人”等思想应用于AI,即难以将AI作为自身投射的实体,这也衍生出关于AI信任、算法厌恶、责任归属等一系列有争议的话题。本文按照“决策要素—决策过程—决策结果”的逻辑顺序展开讨论,结果见表2。

在决策要素方面:AI强大的数据处理能力使其能够有效应对复杂性问题,但AI的决策过程主要依赖于历史数据等显性知识。人类能够获取显性和隐性知识,在处理涉及不确定性和模糊性的问题时具有优势(Boyacı等,2023;Metcalf等,2019)。此外,人类还表现出对创新机会的敏锐嗅觉和创造性思考能力,这使其在决策过程中拥有独特的竞争力。

在决策过程方面:AI在数据处理和决策生成方面展现出快速高效的特质,使其在决策速度和成本方面具备明显优势(Boyacı等,2023;Shrestha等,2019),且AI相对更为公平。人类决策则体现了灵活性、情境判断、同理心和道德(Porayska-Pomsta和Rajendran,2019),但受认知限制(Boyacı等,2023),可能存在明显的偏好。

在决策结果方面:AI在处理大规模数据集时能够实现高准确性的决策(Shrestha等,2019),且非常擅长处理复杂的决策问题(Johnson等,2022)。然而,由于AI决策系统常采用复杂的算法模型,其结果的可解释性通常较差。此外,对于错误决策的责任归属问题也是制约AI普及的重要因素(Porayska-Pomsta和Rajendran,2019)。与此相反,人类通常能够提供相对合理的原因和推理过程,也能够追责到人。但人类存在认知能力的限制,在大数据时代,这可能导致人类难以准确预测未来的发展(Parry等,2016)。

总的来说,AI决策与人类决策各具优势。人类决策者在处理复杂情境和价值判断方面表现出色,而AI在大数据处理、速度、准确性、公平性等方面表现突出。值得注意的是,这些优势并非一成不变。某些情境下,AI决策可能受到算法设计中的偏见影响,导致感知公平性降低。人类决策者也并非始终被视为更友善。因此,如何最大程度地发挥AI和人类决策者的优势,以实现更优决策结果,是一项具有挑战性的任务。

表2 人类决策和AI决策的对比

| | 人类决策 | AI决策 | 参考文献 |
|------|--|--|---|
| 决策要素 | 决策主体 人类 决策类型 不确定性、模糊性决策；具有前瞻性的决策 信息来源 显性知识+隐性知识 决策速度 通常需要时间权衡各方面因素，速度相对较慢 决策成本 受制于有限的认知能力，决策需要的认知成本高 | AI 复杂性决策； 基于历史数据的决策 显性知识 基于给定规则运行，决策速度快 系统处理大规模数据，形成决策成本低 | (Jarrahi, 2018) (Metcalf等, 2019) (Shrestha等, 2019) (Boyacı等, 2023; Shrestha等, 2019) |
| 决策过程 | 决策公平性 容易受主观因素干扰，具有相对明显的个人偏好 决策灵活性 较为灵活、适应性强。可根据情境变化进行调整 决策道德 可以考虑道德、伦理和社会影响等因素 准确性 准确性相对较低 | 给定运作程序，输出结果相对公平、无偏好 需要明确的训练和调整过程，相对固定 相对缺乏同理心和道德判断 准确性相对较高 | (Bigman等, 2021; Langer等, 2023) (Porayska-Pomsta和Rajendran, 2019) (Porayska-Pomsta和Rajendran, 2019) (Shrestha等, 2019) |
| 决策结果 | 可解释性 可解释性相对较高，人类可以描述决策的原因和推理过程 责任归属 有清晰的责任归属 | 可解释性低，内部机制和决策逻辑难以理解和解释 决策出错时难以追究责任 | (Shrestha等, 2019) (Porayska-Pomsta和Rajendran, 2019) |

(二)AI决策与人类决策的互动模式

AI与人类决策关系的相关研究可分为两个方向，即“AI替代人类决策”(替代决策)和“AI与人类协同决策”(合作决策)。替代决策领域的学者认为，随着技术发展，AI能够独立做出此前仅由人类能够完成的决策(Araujo等, 2020)。这些决策不仅涉及日常任务的自动化，同时也扩展至通过学习提供创新的解决方案(Lysyakov和Viswanathan, 2023)。合作模式中，人类和AI共同参与决策过程，共同制定决策。该领域的学者着重强调人机合作的重要性，认为应该将重点从人机竞争转向人机协作，并主张AI系统的设计应旨在增强而不是取代人类(Jarrahi, 2018)。两种决策模式的主要区别见表3。

表3 替代决策和合作决策模式

| 概念 | 替代决策模式 | 合作决策模式 |
|--------|-------------------------------|------------------------|
| 关键思想 | 人类和AI都能够独立形成各自的决策方案，如何选择决策主体？ | 人类和AI如何合作，形成最终的一个决策方案？ |
| 关键特征 | 代理人基本无法干涉AI决策结果 | 代理人影响AI决策的过程或结果 |
| 相关理论 | 聚焦信任和公平的相关理论 | 关注授权方式、信任等各种理论 |
| 主要研究方法 | 实验方法、实证方法 | 实验方法、理论分析、数学建模 |
| 主要研究对象 | 消费者/公众、员工、管理者 | 重点研究决策制定者和管理者 |
| 人类的作用 | 采用者是否引入、消费者或员工表达态度和行为 | 如何和AI共同制定决策方案 |
| 典型场景 | 采购决策、任务分配、贷款决策等 | 医疗、金融投资、组织招聘等 |

四、替代决策模式

“替代”是人类技术向后人类技术转变的核心主题(李河, 2020)。该领域的学者认为，作为决策者的人类本身并不总是不可替代的(Araujo等, 2020)。相关的研究主要关注不同决策主体之间的差异，探讨人们对AI决策与人类自身决策的态度及其影响因素，同时分析这种替代现象可能引发的一系列潜在影响。此类研究可大致分为两类：一是研究哪些因素影响个体对

AI替代决策的接受程度；二是探究AI替代人类决策带来的后果。

基于已有研究,本研究构建了一个替代决策的理论框架(见图5)。该框架表明,当使用AI替代人类进行决策时,个体特征和算法特征可以直接影响个体对替代决策的态度,或通过个体认知间接影响个体,进而导致个体产生抵抗或适应AI的行为。这种行为会对决策有效性产生影响,从而使管理者重新评估是否继续采用AI替代人类决策。此外,随着AI在更重要的决策任务中替代人类,其结果引发了个体、组织和社会层面的广泛讨论。

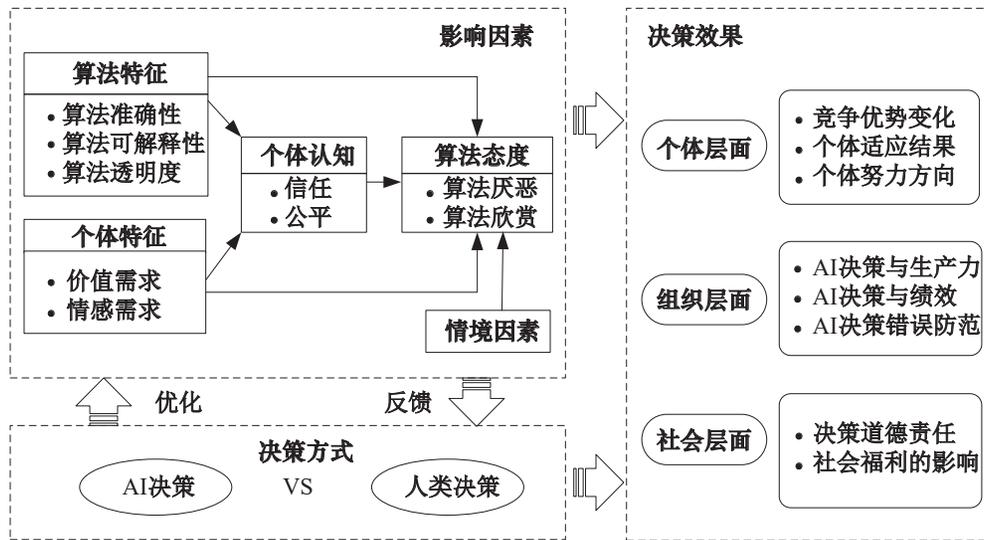


图5 替代决策的理论框架

(一)影响因素

现有研究关于个体对AI决策和人类决策的态度没有形成统一共识,本研究按照个体特征、算法特征和个人认知的分类方式,梳理了每个分类中的关键影响因素。

1.个体特征

价值需求。个体对于自身价值和社会尊重的需求是影响算法态度的重要因素。Kawaguchi (2021)的实验研究发现,员工通常不愿意遵循算法建议,然而,一旦他们的预测被整合到算法中,他们更愿意服从。Castelo等(2023)的研究发现,消费者通常对服务机器人持负面反应,但如果公司能够证明自动化对消费者有益,消费者对服务机器人的反应可能就会等于或优于人类服务提供者。Sharma等(2024)的实证研究发现,努力预期、绩效预期、便利条件和社会影响与客户采用自动化决策呈正相关。Dietvorst和Bartels(2022)的研究结果表明,与人类决策者相比,消费者认为算法更有可能使用最大化(即试图最大化一些测量结果)作为决策策略,这种结果主义决策策略在道德相关的决策中是令人反感的。谢才凤等(2023)指出,认知动机(认知负荷、决策透明)、关系动机(道德地位、人际接触)和存在动机(现实威胁、身份威胁)是个体偏向人类决策而非算法决策的关键因素。

情感需求。人们更倾向于接受人类的建议而不是AI的建议,可能是因为人类可以给予更多的情感支持,而AI缺乏“类人”的情感价值(丁晓军等,2020)。Park等(2022)指出,情绪在家庭成员代表父母拒绝AI决策中起着决定性的作用。Longoni等(2019)发现,独特性忽视是消费者抵制医疗AI的关键因素,消费者认为AI比人类医生更难理解自己的独特特征以及所处的独特环境。Santoro和Monin(2023)指出,随着人们对AI的了解的深入,人类的独特性将被视为更重

要的东西。Kyung和Kwon(2022)的研究表明,个人接受和实施来自AI建议的健康行为的可能性更低,这是因为消费者对AI缺乏情感信任。Hsieh(2023)的研究发现,功能价值、情感价值、个人参与、乐观和创新对医生采用AI辅助诊断的意愿有正向影响。吴继飞等(2020)发现,拟人化能够提升消费者对AI推荐的胜任感,从而提高消费者对AI从事高创造性产品推荐任务的采纳意愿。

2. 算法特征

算法准确性。准确性是指算法在解决问题时能够得到正确或接近真实结果的能力,是评估算法性能的重要指标。个体对AI的态度以及接受基于AI建议的行为意图与感知准确性相关(Chua等,2023)。算法决策质量、透明度以及对数据科学家能力的信任程度越高,人们更容易接受算法建议(Sturm等,2023)。尤其是在高风险决策场景中,人们更倾向于接受先前表现良好的算法的建议(Saragih和Morrison,2022)。

算法可解释性。多数学者强调了AI决策可解释性的重要性(Gaczek等,2023;Kim,2023;Shin,2021;杨炳成等,2023),尤其是对高风险决策(Leichtmann等,2023)。解释会改变用户对可用信息的情境权重,并唤起心智模型的调整(Bauer等,2023)。相比事前解释和视觉解释,事后解释和文本解释的积极效果更明显(Chen等,2024;Ha和Kim,2023)。但也有部分学者质疑是否应该将可解释性作为接受决策的必要条件,甚至质疑是否有必要提供解释。Ghassemi等(2021)认为将可解释性作为AI系统部署的要求可能不现实,因为目前的可解释性方法很难达到预期目标。Westphal等(2023)指出,解释可能损害用户的感知和合法性。Bonezzi等(2022)指出,算法和人类在某种程度上都是“黑匣子”,人们认为他们比算法决策更能理解人类,这可能只是一种错觉。此外,算法可解释性和算法准确性之间需要权衡,复杂算法模型的准确性可能更高,但可解释性会更低(Dwivedi等,2021;Sachan等,2020)。

算法透明度。算法透明度通常被视为可信人工智能的基础(Liu等,2021;邓克涛和张贵红,2024),有助于提高公众信任(樊博和李晶晶,2024)。Grimmelikhuijsen(2023)发现,算法透明度对决策可信度的影响不仅取决于对算法本身的信任,还与其对人类决策者的信任有关。Lehmann等(2022)等的实验研究发现,个体对算法复杂性的感知调节了算法透明度对建议使用的影响,认为算法过于简单可能会严重损害人们对其建议的接受程度。Bankuoru Egala等(2023)发现,算法的不透明、不一致和不准确挑战了临床医生的认知权威,造成了算法厌恶。赵一骏等(2024)发现感知不透明性增加职场中的算法厌恶。

3. 个体认知

缺乏信任是阻碍人们接受AI决策的关键要素(孔祥维等,2022)。人们通常更倾向于信任人类决策者而不是AI决策者,可能是因为AI比人类更有可能做出功利的选择(Zhang等,2022;徐岚等,2024),且AI在决策过程中引发了较少的社交互动,增加了不确定性,降低了用户的信任(Liu,2021)。张语嫣等(2022)将算法拒绝的原因归纳为算法主体怀疑、道德地位缺失和人类特性湮没,分别对应信任、责任和掌控三种心理动机。孙羽佳等(2023)则指出,当前消费者的算法态度正从厌恶逐渐向信任转变。

从信任的细分结构来看,人们对于AI和人类决策的信任来源于不同方面。Bedué和Fritzsche(2022)继承了将信任分为能力、诚信和仁慈三个维度,在组织决策场景中揭示了AI在建立信任方面与传统技术的根本差异。Höddinghaus等(2021)的实验研究发现,AI领导在诚信和透明度方面被认为更高,人类领导者则被认为更具适应性和更加仁慈。Dennis等(2023)的研究发现,作为团队成员的AI被认为具有高能力和诚信,但仁慈程度较低。Langer等(2023)发现参与者对AI决策的初始信任度更低,信任违规和信任修复干预对自动化系统的影响更弱。

Kyung和Kwon(2022)通过实证研究发现,AI缺乏的情感信任(仁慈、正直)而不是认知信任(能力),在接受基于AI的预防性健康干预措施方面起着决定性作用。

公平是影响人们对AI决策态度的另一个关键因素(Ochmann等,2024)。相关研究主要聚焦于AI特征和个体特征如何影响个体对AI决策的公平感知,以及对比个体对AI决策与人类决策公平感知的差异(蒋路远等,2022)。相比人类决策,AI决策通常被认为是更加公平的(Bigman等,2021)。Bai等(2022)的现场实验发现,中国工人认为算法分配过程相比于人工分配更加公平,美国人也认为算法分配更公平。Choung等(2023)通过实验研究发现,求职者认为算法决策比人类做出的决策更公平。然而,也有研究得到了相反结论。裴嘉良等(2021)通过两个实验研究,发现相比上级主管决策,员工认为AI算法决策的信息透明度更低,由此产生更低的程序公平感知。此外,也有研究指出,基于AI的决策可能不仅会使文化刻板印象和歧视永久化,而且会放大这些印象和歧视(Shrestha等,2019)。

针对上述矛盾的结论,一些学者开始探索潜在的边界条件。魏昕等(2021)发现只有面临不利决策时,员工才会觉得算法决策比领导决策更公平,从而展现出更少的组织偏差行为和社会偏差行为。Yu等(2024)发现,由于消费者对机器人服务的期望较低,当消费者被机器人拒绝服务请求时,其对服务的负面评价较低。Efendić等(2020)发现,影响算法信任的关键因素是响应时间,人们认为从算法中缓慢生成的预测不太准确。

4. 个体态度

个体对算法的态度呈现出复杂且多维的特征,主要包括算法厌恶与算法欣赏两大类(张学等,2024)。算法厌恶是人们对于依赖算法决策表现出的不愿意或不信任的一种情感态度(Mahmud等,2022)。尽管AI决策效果往往比人类更好,但人们可能仍然不愿意使用算法(Snijders等,2023; de Véricourt和Gurkan,2023)。Dietvorst和Bharti(2020)通过对4820份调查数据的分析,在投资、医疗等具有内在不确定性的决策领域,证实了该现象。Gill(2020)的研究也指出,尽管自动驾驶汽车的性能更优,但消费者仍然持厌恶态度。Mahmud等(2023)发现,尽管AI具有卓越的能力,但许多管理者仍然反对在决策中使用AI算法。Yan等(2024)通过五项实验研究证实,消费者厌恶基于算法的人力资源决策,无论其决策质量如何。罗映宇等(2023)提炼出算法厌恶的6种成因机制,并阐述了算法厌恶的12种理论边界。

算法欣赏,即人们更愿意接受来自算法的推荐而不是人类的建议,部分研究关注了该现象。Logg等(2019)的六项实验表明,人们更喜欢算法而不是人类决策者。You等(2022)利用实证方法探讨了司法顾问和AI的建议接受情况,发现个体很大程度上表现出算法欣赏。Mahmud等(2022)对80篇经典的实证研究文章进行系统梳理和分析,发现超过50%的研究讨论了算法厌恶,13项研究讨论了算法欣赏,26项研究同时调查了算法厌恶和算法欣赏。

个体的算法态度受到任务特征以及个体特征的共同影响(景怡,2023)。对于更多涉及定量运算的决策,AI被认为更可信,而对于更多涉及情景判断的定性决策,人类的可信任程度更高(冉龙亚等,2022)。如在涉及道德等重要决策(Gill,2020; Waldman和Martin,2022; Zhang等,2022)或决策的不确定程度较高(Dietvorst和Bharti,2020)时,个体往往对算法持有谨慎甚至反感的态度。然而,在任务分配等更为客观的决策领域,个体更加欣赏算法的效率和公平性(Bai等,2022; 王戈和张哲君,2023)。同时,个体自信程度(Snijders等,2023)、感知腐败程度(Castelo,2024)、是否具备专业知识(You等,2022)、归因(高记和冯婧雯,2024),以及对决策任务的熟悉度(Wang,2023)等,都会影响其算法态度。管理者也可能由于传统观念或形象障碍而表现出算法厌恶(Mahmud等,2023)。总之,情境因素对个体算法态度的影响是多维度且复杂的,需要充分考虑这些情境因素,理解个体态度背后的本质原因。

(二)决策效果

1.个体影响

个体层面主要关注AI对个体竞争力的影响以及个体如何应对AI自动化带来的威胁。Krakowski等(2023)发现,AI的采用使人类的传统竞争能力过时,而新的人机能力与传统能力无关甚至是负相关。Tang等(2022)发现,当智能机器具备了自主决策的能力,将与尽职尽责的员工特征重叠,认真的员工不太可能从与智能机器共同工作中获益。Lysyakov和Viswanathan(2023)研究了设计师如何应对引入AI系统的众包设计平台带来的威胁。实证分析发现,引入AI系统后,成功的设计师变得更加专注,并更加注重质量;不成功的设计师提高了竞赛参与度,但并没有改变提交作品的质量。Yam等(2022)发现,提供负面反馈的拟人化机器人主管比非拟人化的机器人更有可能被视为拥有代理权,这导致了更高的人类主管的报复(关闭机器人主管)。员工也会在无法直接与AI交互的情况下会加强和保护其职业角色身份(Strich等,2021)。考虑到AI在常规和标准化任务中的优势,人类面临的选择是放弃这些任务或转向其他任务,并利用创造力、想象力、情感表达和同理心等人类特有的能力与AI展开竞争(Lysyakov和Viswanathan,2023)。

2.组织影响

AI自动化决策通常有助于生产力提升。Luo等(2019)通过对超过6200名客户的实地实验数据分析发现,未公开的聊天机器人在促成客户购买方面与熟练的员工一样有效,比缺乏经验的员工高四倍。Bai等(2022)发现,从算法接收任务可将工人的拣选效率提高15.56%~17.86%。Allal-Chérif等(2021)认为AI能够在不可预测的动态环境中做出复杂的战略决策,为采购部门提供新机会。徐雷等(2024)研究了人工智能算法决策与企业研发“合谋”的情况,指出AI决策下企业研发投入更低而利润水平更高的情况是普遍存在的。值得注意的是,在没有智能的情况下实现自动化可能并不是很有价值(Cui等,2021)。此外,虽然多数情况下AI决策的表现更好,但需要与组织实际相结合。如Li等(2021)发现自动化的订购系统可为零售商带来更高的利润,但在去中心化供应链中可能适得其反。Harada等(2021)发现,具有AI驱动差异诊断列表和不具有的组之间没有诊断准确率的差异。

AI决策也可能产生意外的后果(Mayer等,2020)。一方面,消费者在与AI互动时跟和人类互动时的道德关注和行为是不同的,AI的使用可能提高消费者从事不道德行为的可能性(Kim等,2023),并影响人际关系(Claire等,2023)。根据Giroux等(2022)的研究,AI结账不太可能出现道德意图(报告错误的意图),但当机器更像人时,道德意图就会增加。Zhou等(2022)发现人—AI互动通过支持功利主义判断对参与者的捐赠产生了负面影响。另一方面,AI的有效实施和管理是潜力发挥的关键。针对美国医院的实地调研发现,高性能的AI工具在实际应用中并未达到预期,这是因为医生在处理不确定性问题时通常依靠自己的专业知识,而这些知识并未被纳入AI工具中(Lebovitz等,2021)。此外,错误的建议可能导致重大失误(Dwivedi等,2021),接受AI建议可以降低责任风险,但也容易诱导投机行为。同时,也可能会增加组织的近视学习(Balasubramanian等,2022),降低人类的批判性思维能力(Budhwar等,2023)。

3.社会影响

AI决策也引发了人们对道德和社会福利的关注(Shrestha等,2019)。自动驾驶汽车发展面临的核心伦理困境是:是否应该将道德规范嵌入算法结构以及应当以何种方式嵌入(Young和Monroe,2019;白惠仁,2019)。更严重的是,AI可能改变人们的道德判断。Ghasemaghaci和Kordzadeh(2024)在一项涉及122名管理人员的实验研究中,发现算法不公正会导致歧视性决策,但不会增加内疚感。Gill(2020)发现,人们更容易接受自动驾驶汽车对行人造成的伤害,这

种道德判断的改变是由责任转移驱动的。此外,部分学者担心当决策权力转移到机器手中时,人类会失去对道德决策的控制(Moser等,2022)。

在社会福利问题方面,已有研究指出,相比要求待遇对等的现行法律,要求影响对等的机器学习算法可能会降低更准确算法对企业的好处,这可能导致以利润最大化为目标的企业在机器学习方面投资不足(Fu等,2022)。此外,决策的真正来源可能会被人为混淆,从而导致责任转移。如何落实算法责任议题与构建可持续性算法商业模式,确保AI的应用对组织和社会福利产生积极影响,也是相关领域的重要话题(肖红军,2022)。

五、合作决策模式

新兴研究开始探索如何通过AI提升人类能力而不是取代人类(Anthony等,2023; Metcalfe等,2019)。这种重视人机合作、实现协同决策的观点反映了对于AI决策的新思考。相关研究集中在两个方面:一是通过定性分析揭示人机合作形成决策的潜在模式和操作路径;二是通过实验等定量方法研究AI对合作决策的影响及合作效果的优化。

(一)合作机制设计

AI的崛起对组织决策的设计和构建带来了新的挑战,通过合理设计和构建人与AI共同参与的决策模式,可以充分发挥人机协同的优势,从而实现更高效、准确和创新的组织决策。为此,学者们提出了多种合作决策设计方案。其中之一便是将AI作为决策支持工具,为人类决策者提供数据和分析结果,从而提升决策质量和多样性。例如,Loftus等(2020)的研究发现,将AI与手术决策集成,可以提供更准确全面的数据和分析结果,帮助医生更好地理解手术决策的影响,从而改善现有的护理方式。

另一种设计方案是AI与人类决策者协同决策,通过充分发挥双方优势,形成更高质量的决策。Jarrahi(2018)指出结合AI的信息收集与分析速度以及人类卓越的直觉和洞察力,可以更好地应对决策中的复杂性、不确定性和模糊性。Terziyan等(2018)指出,借鉴人类(直觉、情感、非理性等)决策的经典实践,可以促进AI系统的改进,更好地模仿和运用人类在决策过程中的优势。Vincent(2021)提出了一种整合了直觉和AI优势的决策模型,但也指出只有当任务结构不合理且决策者是领域专家时,直觉和AI的结合才会有效。

还有一些学者提出了协同智能的决策设计方案。通过将决策任务分配给多个AI系统和人类决策者,利用集体智慧来达成决策目标。Metcalfe等(2019)提出组织可以利用人工群智能(ASI)的协作技术来增强人类群体的智能,促进更好的商业决策。Spring等(2022)通过对两家律师事务所和两家会计师事务所的多案例研究,发现AI系统在专业服务流程的三个阶段(诊断、推理和治疗)中被有选择地使用。Choudhary等(2023)提出了一种专业化分工的替代方案,使人类和AI处理相同的问题,将结果以某种方式(如估计平均值、分类问题的法定人数、多数或一致)聚合,得出最终结果。

(二)AI对合作决策的影响与合作效果优化

合作领域的应用研究主要聚焦于AI在合作决策中的影响及合作效果的优化,包括四个部分。具体内容见图6。

1.AI对决策过程的影响

关于AI对人类决策过程的影响,现有研究强调了元知识的关键作用。Jussupow等(2021)首次提出医生在评估AI建议时,使用元认知^①来监测和控制他们的推理。他们认为错误的诊断决

^①元认知(metacognition)是指监控和控制人类决策的二阶认知过程。在决策过程中,不同类型的元认知帮助决策者反思自己的推理,决定考虑哪些信息以及如何解释这些信息。

策往往是医生在利用与自己的推理相关的元认知,而在与AI相关联的元认知方面存在缺陷。Fügenger等(2022)指出,由于缺乏元知识^①,从根本上限制了人类决策者与AI的协作。除了决策者本身,现有研究还关注了管理者态度。Estep等(2024)的实验研究发现,只有财务主管所在的公司使用AI时,财务主管才会对基于AI的审计结果做出更大的调整。Keding和Meissner(2021)发现在研发投资环境中,高管们比人类顾问更倾向于遵循基于AI的建议。

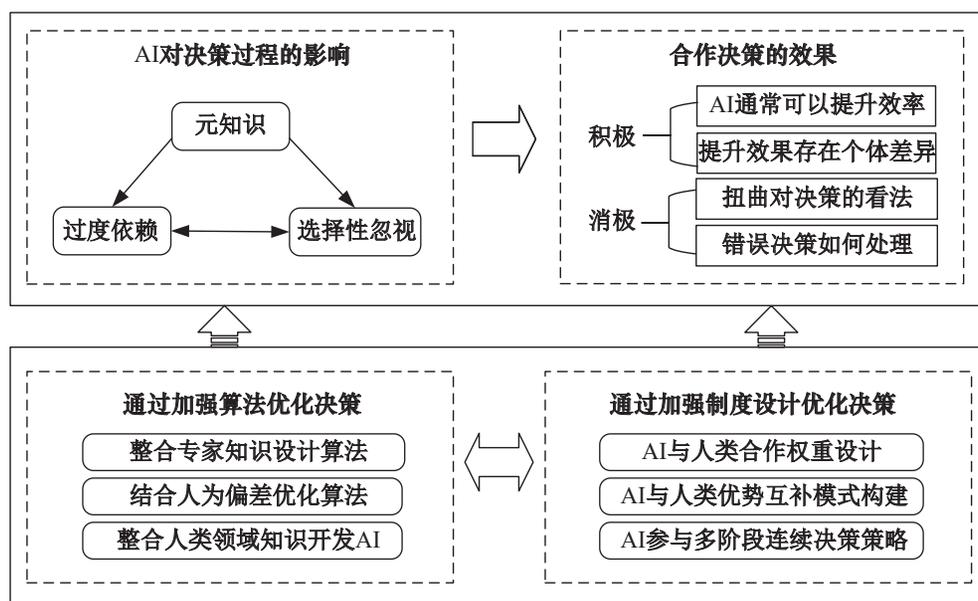


图6 AI与人类合作决策的主要研究内容

信任是人机成功合作的基础,个体在人机交互中可能会出现信任偏差:过度信任和信任不足(黄心语和李晔,2024),表现为“过度依赖”或“选择性忽视”AI决策的倾向(Alon-Barkat和Busuioac,2023;Neumann等,2023)。前者可能导致人类决策者的重要性降低,甚至使人类与AI合作决策名存实亡,后者则违背了引入AI决策的初衷。Lebovitz等(2022)等通过实地实验研究了三个放射科部门的医生使用AI的情况,发现只有一个部门的医生能够将AI结果与他们最初的判断联系起来,另外两个部门的医生要么经常忽略AI的建议,要么在没有太多反思的情况下接受AI的建议。这可能是由于决策者对AI决策的内在不信任引起的,也可能是因为决策者在任务环境中难以准确判断AI决策是否比自己做出的决策更好或更差(de Véricourt和Gurkan,2023)。目前尚不清楚组织管理者是否会从AI使用中受益,还是成为只会宣布AI决策结果的“傀儡”(Leyer和Schneider,2021)。

2.AI合作决策效果

人机协同决策通常优于AI和人类单独决策(Boyacı等,2023)。这种合作决策效果的提升是基于决策者愿意考虑AI的决策建议,并依赖于二者的协同。Shin等(2023)发现超级AI的发展促使人类玩家摆脱传统策略,并诱导他们探索新动作,这反过来改善了他们的决策。个体对AI决策的态度以及决策提升效果存在异质性。Gaube等(2021)发现,放射科医生认为来自AI系统的建议质量较低,而任务专业知识较少的医生则持相反意见。Tschandl等(2020)针对临床决策的实验研究发现,与AI或单独的医生相比,高质量的基于AI的临床决策支持可以提高诊断的准确性,而经验最少的临床医生从基于AI的支持中获益最多。

^①元知识(metaknowledge)是指“关于知识的知识”,这里将元知识定义为评估自身能力的的能力。

尽管协同决策具有优势,但目前协同决策难以发挥应有的价值。首先,AI可能被当作证实想法的工具。Ahmad等(2023)发现警察只遵循与他们判断一致的AI建议。其次,AI决策可能会扭曲决策者对决策结果的看法。Keding和Meissner(2021)利用150名高级管理人员的实验数据发现,基于AI的建议增加了管理者采取投资的可能性,并提高了感知的决策质量。此外,AI提供错误的决策可能会严重影响决策质量。当参与者获得不正确的算法支持时,特别是当他们在提供自己的判断之前获得算法支持时,准确性降低(Agudo等,2024)。Tschantl等(2020)发现错误的AI决策会误导包括专家在内的所有临床医生。Gaube等(2021)的研究发现,当参与者收到不准确的建议时,无论声称的来源是什么,诊断的准确性都会显著降低。

3.考虑人类特征的合作决策优化

设计合理的决策机制可以有效提高决策质量。其中一种方法是通过整合专家知识来改进决策算法。如Terziyan等(2018)等提出了Pi-Mind技术,该技术能够获取特定人类专家的基本参数(包括直觉和情感等),实现自主和快速决策,并让人类和AI共同承担决策后果的责任。另一种方法是通过考虑人为偏差来优化算法。如Sun等(2022)等提出了以人为中心的垃圾箱包装算法,该算法预测并纳入人类行为偏差,从而提高包装效率和时间利用率。Liu-Thompkins等(2022)将人工智能同理心,即AI情感共鸣的能力,编码至AI算法及代理中,进而融入AI支持的营销活动中。此外,为了解决战略决策问题,Wu等(2023)构建了一个系统的决策框架。Gonzalez(2023)提出了一种通用的认知算法,模拟动态环境中的人类决策。

4.考虑人机协作机制设计的决策优化

该领域重点关注如何设计人与AI的合作机制,从而更好地提高决策质量(Boyacı等,2023; Haesevoets等,2021)。一些学者试图通过合理构建AI与人类的合作权重优化决策。Costello等(2020)发现,在借贷决策中加入人类的自由裁决权可以提高贷款结果。Haesevoets等(2021)对1025名经理的调查数据表明,大多数管理者愿意接受与机器的合作伙伴关系,只要他们觉得决策主要是根据人类的投入和判断做出的。另一些学者则基于优势互补原则,致力于构建高效的人机协同决策模式。Boyacı等(2023)考虑了人类的灵活性和有限的认知能力,以及机器强大的数据处理能力,建立了对应的数字模型,结果发现,机器输入总体上能提高人类决策的准确性。还有学者关注了多阶段连续协同决策,Wesche和Sonderegger(2021)发现AI决策的效果因招聘阶段而有所不同,面试阶段使用AI对申请人感知到的组织吸引力和申请意愿产生了负面影响,筛选阶段的负面评价较小。Langer等(2021)发现在处理申请者信息之前通过AI获得申请者排名时,参与者的自我效能感更高;在处理申请者信息之后获得排名时,参与者的决策满意度较高。

六、AI赋能下的综合决策框架

替代决策和合作决策两种模式并不相互排斥,二者相辅相成,共同推动AI在组织决策领域的发展(Einola和Khoreva,2023)。然而,在特定情境下,一种模式可能会占据主导地位。本文开发了综合的决策框架来整合两个研究领域,以期提供更系统的见解。开发综合框架的关键是观察到两个研究领域存在重叠和互补。虽然替代和合作决策基于不同的假设,在微观决策过程中存在根本性差别,但从更广泛的视角来看,AI自动化决策也需要人的参与,包括接受、拒绝和修订算法反馈等。过度强调AI强化人类决策或AI自动化决策会助长强化循环,带来负面的组织和社会后果(Raisch和Krakowski,2021)。

按照“决策模式选择—决策过程—决策结果”的顺序,构建如图7所示的AI赋能下的综合决策框架。针对特定领域的决策任务,组织需要选择适当的决策模式。这种选择受到多种因素的影响,包括管理者的态度、决策者的偏好以及外部环境的变化。在决策过程中,AI和人类作为

两个决策主体,形成了两种不同的决策互动模式,即替代决策和合作决策。替代决策强调人或AI独立作为主导决策者,而合作决策则注重人类与AI之间的协同合作。二者都受到AI和人类自身特征、能力及决策任务性质的影响。不同的是,替代决策领域的研究主要关注特定情境下人和AI作为不同决策主体的影响,合作决策领域则倾向于考虑合作机制的设计及AI决策对人类认知的潜在影响。决策结果方面,替代决策和合作决策两个领域的研究都关注利益相关者态度、行为和决策价值的评估,前者更关注利益相关者的反应,后者则更注重决策价值的评估,最终结果体现在企业利润和社会福利方面。

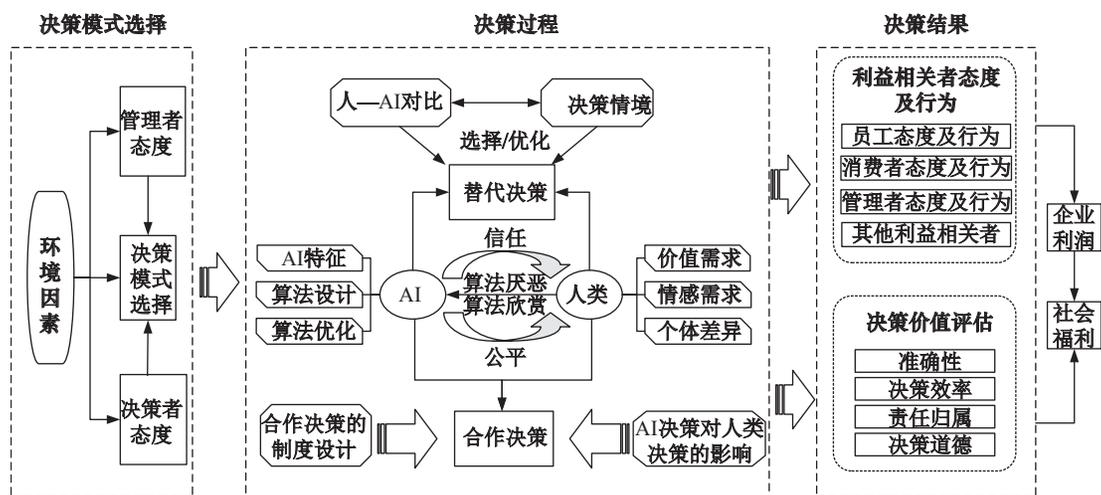


图7 AI赋能下的综合决策框架

从综合决策框架可以看出,替代决策和合作决策领域的研究既有相似性,也存在一定的区别。个体对AI决策的态度是影响决策模式选择的关键因素。在组织中,管理者和决策者对AI决策的看法,直接决定了AI决策能否被采纳和推广。决策效率是替代决策和合作决策领域共同关注的问题。有效的算法管理是提高决策效率的关键途径,包括提升AI的实际性能,使其在组织内部得到有效应用,以及优化AI算法与人类决策者之间的信息传递和沟通协作。通过加强对算法的设计和优化,可以发挥AI算法在决策过程中的优势,实现决策的高效性和准确性。不同的是,替代决策和合作决策在决策过程和决策结果方面存在差异。合作决策需要考虑更多的因素,包括制度设计、AI算法对决策者的潜在影响等。与此同时,由于替代决策和合作决策领域的研究都处于探索阶段,目前两个领域在决策结果部分存在重叠,但侧重点有所不同。

在两种决策方式得以完整实施的过程中,需要注意以下关键因素:(1)管理者和决策者态度。管理者和决策者对于使用AI决策系统的态度是决策方式的基本前提。他们的态度受到个体、环境、任务、组织和文化等多重因素的影响,并且他们之间的态度可能存在相互影响。(2)管理者需要制定合适的算法管理策略。在使用AI算法进行决策时,可以通过优化和管理算法来改进决策质量。在合作决策过程中,人类决策者可能受到AI决策结果的干扰,从而产生有偏差的决策。(3)合作决策的制度设计需要结合具体情境形成定制化方案。这种方案的设计依赖于科学的理论设计,需要充分考虑组织环境、任务特征和组织文化,结合实际决策场景进行探索。(4)利益相关者的态度和行为。除了考虑决策质量,管理者还需要考虑与决策密切相关的利益相关者(如员工、消费者等)的态度和行为。在追求效率的同时注重“以人为本”的原则。尤其是在涉及有重大影响的公众决策中,政府需要慎重制定合理的激励政策,以平衡企业利润和社会福利。

七、结论及研究展望

本文通过系统回顾组织管理领域的AI决策相关文献,首次从决策要素、决策过程和决策结果三个维度剖析了AI决策与人类决策的本质差异。同时理清了AI决策与人类决策的两种互动模式,并明确了各自的研究重点。在此基础上,按照“决策模式选择—决策过程—决策结果”的分析思路,构建了AI赋能下的综合决策框架。

(一)探究AI决策的影响因素及效果

1.探究高层管理者和决策者对AI决策的态度如何相互影响。现有研究多关注管理者和决策者各自对AI决策的态度,尤其重视决策者的态度,但很少探讨二者之间的相互影响(Haesevoets等,2021;Liang和Xue,2022)。高层管理者对AI决策的战略性判断和决策者对AI决策的战术性判断可能存在偏差(Estep等,2024),尤其涉及社会地位、代理权利等复杂认知因素时(Yam等,2022)。当存在这种态度差异时,AI决策的结果可能更加复杂多变(Mayer等,2020)。未来可进一步探讨管理者和决策者态度的相互影响及其转变。此外,有必要持续关注AI参与决策对组织领导者和员工心理层面的影响(张广胜和杨春荻,2020)。

2.增强对AI与人类合作决策效果的关注。在替代决策领域,已有研究比较了AI决策和人类决策下个体态度和行为的影响(裴嘉良等,2021;魏昕等,2021)。通过揭示关键的作用机制,并尝试通过调整管理措施来改善负面态度和消极行为(Kyung和Kwon,2022;Sun等,2022)。然而,在合作决策领域,由于缺乏统一的决策框架机制,限制了对这种决策方式结果的讨论。未来,随着对合作决策的深入探讨,人们将更加关注这些决策中机器行为和人的行为的交互机理和相互影响,尤其是机器行为对人的塑造和影响(曾大军等,2021b)。

3.关注复杂决策场景下,替代决策和合作决策共存可能带来的影响。这两种决策模式可以相互补充,主要取决于具体的决策场景和决策问题的复杂程度。如果涉及团队决策,AI可能替代部分决策者,这也意味着剩余的其他人需要与AI协同决策。在这类决策情境中,替代决策的发生可能会影响剩余决策者对与AI合作决策的态度和行为。从更微观的视角出发,深入挖掘这种现象蕴含的底层逻辑,有助于更好地理解和洞察AI参与决策的深层次作用效果。

4.现有相关领域的定量研究多采用实验研究方法,未来研究可以考虑多种研究方法的综合应用。比如通过QCA来检查不同场景下多条件组态,以深入理解采用AI决策的复杂因果关系(Cao等,2021);使用仿真方法研究有AI参与的多主体复杂互动决策的形成和演化过程;还可以通过决策主体智能建模来理解复杂环境下的决策行为,深入洞察多个参与主体之间的互动关系以及这些关系如何影响最终的决策结果(曾大军等,2021a)。

(二)设计有效的人—AI决策机制

1.在AI和人类均可独立决策情境下,探究是否以及多大程度上为人类决策者设置否决权。设置否决权可以确保人类对决策结果有一定的控制权,但过度的否决权可能会削弱AI决策的准确性和公平性等优势(Boyaci等,2023;Costello等,2020)。如何根据任务场景、决策风险等的不同,设计有利于组织利益和社会福利的否决机制,这些权利如何随时间推移演变(Costello等,2020)?如何确保AI决策系统不被报复性关闭(Yam等,2022)?此外,随着AI越来越多地执行具有潜在危险并可能对人造成伤害的功能,如何有效管理由替代决策造成的重大错误,采用何种归责方式,以及如何进行补救(Dwivedi等,2021)?这些都是组织在重要决策中是否允许AI替代人类决策需要考虑的重要问题。

2.设计有效的人与AI协同决策制度。如何将前沿算法技术有效嵌入决策过程、真正实现决策质量提升还需解析一系列人机协同决策中的基础机制(刘伦,2023)。未来可以考虑:(1)如何

确定决策顺序?混合顺序决策可以充分发挥人类和AI决策的优势,但也可能放大彼此的弱点(Shrestha等,2019)。(2)如何确定决策权重?在重要决策场景中,分配给AI正确的决策权重是一个具有挑战性的问题(Parry等,2016)。(3)决策如何随时间动态变化?人类和AI也是在相互学习中不断成长的,人类可以根据过往使用AI的历史经验改变其对AI决策的看法,从而更倾向于接受或拒绝AI决策。决策者与AI之间的责任、权力如何随时间推移而变化也是未来的重要研究课题(Fast和Schroeder,2020;Haesevoets等,2021)。

3.针对团队层面AI决策机制的深入探索仍显匮乏(Dennis等,2023)。团队决策是一个涉及多主体、多因素、多层次的复杂动态过程。AI的引入为组织决策提供了新工具,但同时也带来了新挑战(Mayer等,2020)。虽然AI可能提高个人和较小群体的表现,但可能损害较大群体的表现(Fügener等,2021)。团队层面的AI决策研究不仅要关注AI技术本身的发展与应用,更要深入探讨AI如何与团队成员形成有效互动、如何影响团队决策的流程与结果,以及如何优化团队的整体绩效。对AI的使用及其对决策的影响进行理论化,构建综合的概念框架,是系统地理解AI决策的关键所在(Duan等,2019)。鉴于AI在决策中的角色尚未清晰,有必要构建包含个体、团队和组织的多层次、多维度的AI决策研究框架,以全面理解AI在决策中的复杂作用。

(三)聚集AI决策的本土化研究

AI决策的本土化研究尤为重要。当前,国内学者已经注意到AI决策在透明度、可解释性、公平性等方面的问题(邓克涛和张贵红,2024),并讨论了人们的算法态度(罗映宇等,2023;孙羽佳等,2023)。少部分定量研究探索了员工对不同决策主体的感知公平性(裴嘉良等,2021;魏昕等,2021),以及算法可解释性对消费者采纳意愿的影响(杨炳成等,2023)。尽管国内相关领域的研究处于起步阶段,但我国在AI决策的研究方面存在显著优势。除了有利的政策和资金支持,国内还拥有大量的数据资源和丰富的AI决策场景,为AI决策研究提供了便捷的条件。

为了优化AI决策与人类决策之间的关系并推进其在中国企业的有效应用,未来可以从接受度、文化融合以及实际场景应用等多个维度进行深入探究。首先,要重点关注中国情境下AI决策的接受度及其影响因素。中国拥有独特的文化背景和社会结构,需要持续关注个体差异、算法特征的影响及其他可能的边界条件,以更好地适应中国市场的需求。其次,AI决策机制的设计需要充分考虑本土文化和价值观的特点,探究AI与人类的交互及协同决策过程以及实现彼此增强的边界条件,构建AI决策与本土文化和价值观有效融合的组织决策系统。最后,基于替代和合作决策互补的实际场景开展研究,探究不同决策场景下AI与人类决策者的互动方式,以及这种互动如何影响组织的决策质量和效率。

主要参考文献

- [1]高记,冯婧雯.不利结果下算法决策对员工公平感的双刃剑效应:基于归因理论的视角[J].中国人力资源开发,2024,41(5):36-53.
- [2]蒋路远,曹李梅,秦昕,等.人工智能决策的公平感知[J].心理科学进展,2022,30(5):1078-1092.
- [3]景怡.算法态度文献综述、研究框架和未来展望[J].华东经济管理,2023,37(5):107-118.
- [4]孔祥维,唐鑫泽,王子明.人工智能决策可解释性的研究综述[J].系统工程理论与实践,2021,41(2):524-536.
- [5]罗映宇,朱国玮,钱无忌,等.人工智能时代的算法厌恶:研究框架与未来展望[J].管理世界,2023,39(10):205-227.
- [6]裴嘉良,刘善仕,钟楚燕,等.AI算法决策能提高员工的程序公平感知吗[J].外国经济与管理,2021,43(11):41-55.
- [7]孙羽佳,苏淞,唐红红.基于信任视角的消费者算法态度研究述评与展望[J].经济管理,2023,45(10):188-208.
- [8]魏昕,黄鸣鹏,李欣悦.算法决策、员工公平感与偏差行为:决策有利性的调节作用[J].外国经济与管理,2021,43(11):56-69.
- [9]谢才凤,邬家骅,许丽颖,等.算法决策趋避的过程动机理论[J].心理科学进展,2023,31(1):60-77.

- [10]曾大军, 张柱, 梁嘉琦, 等. 机器行为与人机协同决策理论和方法[J]. *管理科学*, 2021, b,34(6): 55-59.
- [11]Alon-Barkat S, Busuioc M. Human-AI interactions in public sector decision making: “Automation bias” and “selective adherence” to algorithmic advice[J]. *Journal of Public Administration Research and Theory*, 2023, 33(1): 153-169.
- [12]Araujo T, Helberger N, Kruike-meier S, et al. In AI we trust? Perceptions about automated decision-making by artificial intelligence[J]. *AI & Society*, 2020, 35: 611-623.
- [13]Balasubramanian N, Ye Y, Xu M T. Substituting human decision-making with machine learning: Implications for organizational learning[J]. *Academy of Management Review*, 2022, 47(3): 448-465.
- [14]Bigman Y E, Yam K C, Marciano D, et al. Threat of racial and economic inequality increases preference for algorithm decision-making[J]. *Computers in Human Behavior*, 2021, 122: 106859.
- [15]Bonezzi A, Ostinelli M, Melzner J. The human black-box: The illusion of understanding human better than algorithmic decision-making[J]. *Journal of Experimental Psychology: General*, 2022, 151(9): 2250-2258.
- [16]Boyacı T, Canyakmaz C, de Véricourt F. Human and machine: The impact of machine input on decision making under cognitive limitations[J]. *Management Science*, 2023, 70(2): 1258-1275.
- [17]Cao G M, Duan Y Q, Edwards J S, et al. Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making[J]. *Technovation*, 2021, 106: 102312.
- [18]Costello A M, Down A K, Mehta M N. Machine + man: A field experiment on the role of discretion in augmenting AI-based lending models[J]. *Journal of Accounting and Economics*, 2020, 70(2-3): 101360.
- [19]Dietvorst B J, Bartels D M. Consumers object to algorithms making morally relevant tradeoffs because of algorithms' consequentialist decision strategies[J]. *Journal of Consumer Psychology*, 2022, 32(3): 406-424.
- [20]Duan Y Q, Edwards J S, Dwivedi Y K. Artificial intelligence for decision making in the era of big data – evolution, challenges and research agenda[J]. *International Journal of Information Management*, 2019, 48: 63-71.
- [21]Dwivedi Y K, Hughes L, Ismagilova E, et al. Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy[J]. *International journal of information management*, 2021, 57: 101994.
- [22]Fu R S, Aseri M, Singh P V, et al. “Un” fair machine learning algorithms[J]. *Management Science*, 2022, 68(6): 4173-4195.
- [23]Fügener A, Grahl J, Gupta A, et al. Cognitive challenges in human-artificial intelligence collaboration: Investigating the path toward productive delegation[J]. *Information Systems Research*, 2022, 33(2): 678-696.
- [24]Grimmelikhuijsen S. Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making[J]. *Public Administration Review*, 2023, 83(2): 241-262.
- [25]Haesevoets T, De Cremer D, Dierckx K, et al. Human-machine collaboration in managerial decision making[J]. *Computers in Human Behavior*, 2021, 119: 106730.
- [26]Höddinghaus M, Sondern D, Hertel G. The automation of leadership functions: Would people trust decision algorithms?[J]. *Computers in Human Behavior*, 2021, 116: 106635.
- [27]Jarrahi M H. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making[J]. *Business Horizons*, 2018, 61(4): 577-586.
- [28]Jussupow E, Spohrer K, Heinzl A, et al. Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with artificial intelligence[J]. *Information Systems Research*, 2021, 32(3): 713-735.
- [29]Kawaguchi K. When will workers follow an algorithm? A field experiment with a retail business[J]. *Management Science*, 2021, 67(3): 1670-1695.
- [30]Keding C, Meissner P. Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions[J]. *Technological Forecasting and Social Change*, 2021, 171: 120970.
- [31]Langer M, König C J, Back C, et al. Trust in artificial intelligence: Comparing trust processes between human and automated trustees in light of unfair bias[J]. *Journal of Business and Psychology*, 2023, 38(3): 493-508.
- [32]Langer M, Landers R N. The future of artificial intelligence at work: A review on effects of decision automation and

- augmentation on workers targeted by algorithms and third-party observers[J]. *Computers in Human Behavior*, 2021, 123: 106878.
- [33]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[J]. *Organization Science*, 2022, 33(1): 126-148.
- [34]Lehmann C A, Haubitz C B, Fügener A, et al. The risk of algorithm transparency: How algorithm complexity drives the effects on the use of advice[J]. *Production and Operations Management*, 2022, 31(9): 3419-3434.
- [35]Leyer M, Schneider S. Decision augmentation and automation with artificial intelligence: Threat or opportunity for managers?[J]. *Business Horizons*, 2021, 64(5): 711-724.
- [36]Liang H G, Xue Y J. Save face or save life: Physicians' dilemma in using clinical decision support systems[J]. *Information Systems Research*, 2022, 33(2): 737-758.
- [37]Logg J M, Minson J A, Moore D A. Algorithm appreciation: People prefer algorithmic to human judgment[J]. *Organizational Behavior and Human Decision Processes*, 2019, 151: 90-103.
- [38]Longoni C, Bonezzi A, Morewedge C K. Resistance to medical artificial intelligence[J]. *Journal of Consumer Research*, 2019, 46(4): 629-650.
- [39]Luo X M, Tong S L, Fang Z, et al. Frontiers: Machines vs. Humans: The impact of artificial intelligence chatbot disclosure on customer purchases[J]. *Marketing Science*, 2019, 38(6): 913-1084.
- [40]Mahmud H, Islam A K M N, Ahmed S I, et al. What influences algorithmic decision-making? A systematic literature review on algorithm aversion[J]. *Technological Forecasting and Social Change*, 2022, 175: 121390.
- [41]Mahmud H, Islam A K M N, Mitra R K. What drives managers towards algorithm aversion and how to overcome it? Mitigating the impact of innovation resistance through technology readiness[J]. *Technological Forecasting and Social Change*, 2023, 193: 122641.
- [42]Mayer A S, Strich F, Fiedler M. Unintended consequences of introducing AI systems for decision making[J]. *Mis Quarterly Executive*, 2020, 19(4): 6.
- [43]Metcalf L, Askay D A, Rosenberg L B. Keeping humans in the loop: Pooling knowledge through artificial swarm intelligence to improve business decision making[J]. *California Management Review*, 2019, 61(4): 84-109.
- [44]Park E H, Werder K, Cao L, et al. Why do family members reject AI in health care? Competing effects of emotions[J]. *Journal of Management Information Systems*, 2022, 39(3): 765-792.
- [45]Pietronudo M C, Croidieu G, Schiavone F. A solution looking for problems? A systematic literature review of the rationalizing influence of artificial intelligence on decision-making in innovation management[J]. *Technological Forecasting and Social Change*, 2022, 182: 121828.
- [46]Raisch S, Krakowski S. Artificial intelligence and management: The automation-augmentation paradox[J]. *Academy of Management Review*, 2021, 46(1): 192-210.
- [47]Saragih M, Morrison B W. The effect of past algorithmic performance and decision significance on algorithmic advice acceptance[J]. *International Journal of Human-Computer Interaction*, 2022, 38(13): 1228-1237.
- [48]Sharma S, Islam N, Singh G, et al. Why do retail customers adopt artificial intelligence (AI) based autonomous decision-making systems?[J]. *IEEE Transactions on Engineering Management*, 2024, 71: 1846-1861.
- [49]Shrestha Y R, Ben-Menahem S M, von Krogh G. Organizational decision-making structures in the age of artificial intelligence[J]. *California Management Review*, 2019, 61(4): 66-83.
- [50]Strich F, Mayer A S, Fiedler M. What do I do in a world of artificial intelligence? Investigating the impact of substitutive decision-making AI systems on employees' professional role identity[J]. *Journal of the Association for Information Systems*, 2021, 22(2): 304-324.
- [51]Sturm T, Pumplun L, Gerlach J P, et al. Machine learning advice in managerial decision-making: The overlooked role of decision makers' advice utilization[J]. *The Journal of Strategic Information Systems*, 2023, 32(4): 101790.
- [52]Sun J K, Zhang D J, Hu H Y, et al. Predicting human discretion to adjust algorithmic prescription: A large-scale field experiment in warehouse operations[J]. *Management Science*, 2022, 68(2): 846-865.

- [53]Syam N, Sharma A. Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice[J]. *Industrial Marketing Management*, 2018, 69: 135-146.
- [54]Tang P M, Koopman J, McClean S T, et al. When conscientious employees meet intelligent machines: An integrative approach inspired by complementarity theory and role theory[J]. *Academy of Management Journal*, 2022, 65(3): 1019-1054.
- [55]Terziyan V, Gryshko S, Golovianko M. Patented intelligence: Cloning human decision models for industry 4.0[J]. *Journal of Manufacturing Systems*, 2018, 48: 204-217.
- [56]Tschandl P, Rinner C, Apalla Z, et al. Human-computer collaboration for skin cancer recognition[J]. *Nature Medicine*, 2020, 26(8): 1229-1234.
- [57]Vincent V U. Integrating intuition and artificial intelligence in organizational decision-making[J]. *Business Horizons*, 2021, 64(4): 425-438.
- [58]Wesche J S, Sonderegger A. Repelled at first sight? Expectations and intentions of job-seekers reading about AI selection in job advertisements[J]. *Computers in Human Behavior*, 2021, 125: 106931.
- [59]Westphal M, Vössing M, Satzger G, et al. Decision control and explanations in human-AI collaboration: Improving user perceptions and compliance[J]. *Computers in Human Behavior*, 2023, 144: 107714.
- [60]Wu C, Zhang R, Kotagiri R, et al. Strategic decisions: Survey, taxonomy, and future directions from artificial intelligence perspective[J]. *ACM Computing Surveys*, 2023, 55(12): 250.
- [61]Yam K C, Goh E Y, Fehr R, et al. When your boss is a robot: Workers are more spiteful to robot supervisors that seem more human[J]. *Journal of Experimental Social Psychology*, 2022, 102: 104360.
- [62]Yan C F, Chen Q, Zhou X Y, et al. When the automated fire backfires: The adoption of algorithm-based hr decision-making could induce consumer's unfavorable ethicality inferences of the company[J]. *Journal of Business Ethics*, 2024, 190(4): 841-859.
- [63]You S, Yang C L, Li X T. Algorithmic versus human advice: Does presenting prediction performance matter for algorithm appreciation?[J]. *Journal of Management Information Systems*, 2022, 39(2): 336-365.
- [64]Young A D, Monroe A E. Autonomous morals: Inferences of mind predict acceptance of AI behavior in sacrificial moral dilemmas[J]. *Journal of Experimental Social Psychology*, 2019, 85: 103870.
- [65]Yu S B, Xiong J, Shen H. The rise of chatbots: The effect of using chatbot agents on consumers' responses to request rejection[J]. *Journal of Consumer Psychology*, 2024, 34(1): 35-48.
- [66]Zhang Z X, Chen Z S, Xu L Y. Artificial intelligence and moral dilemmas: Perception of ethical decision-making in AI[J]. *Journal of Experimental Social Psychology*, 2022, 101: 104327.
- [67]Zhou Y Y, Fei Z Y, He Y Q, et al. How human-chatbot interaction impairs charitable giving: The role of moral judgment[J]. *Journal of Business Ethics*, 2022, 178(3): 849-865.

AI Decision-making in Organizational Management: A Review and Prospects

Zhang Yali, Li Liaoliao, Ding Zhenbin

(School of Management, Northwestern Polytechnical University, Xi'an 710129, China)

Summary: This paper focuses on AI decision-making in the field of organizational management, systematically reviewing 176 domestic and foreign literatures published from 2018 to 2024, and analyzing the complex interactions between AI and humans in organizational decision-making. The results reveal the essential differences between AI decision-making and human decision-making from three dimensions: decision-making elements, decision-making processes, and decision-making

outcomes, and extracts two interaction modes between AI and human decision-making: substitutive decision-making and collaborative decision-making. The research on substitutive decision-making centers on the disparities among decision-making entities and their potential consequences; while the research on collaborative decision-making emphasizes the design of collaborative mechanisms, the impact of AI on collaborative processes, and the optimization of collaborative outcomes. On this basis, a comprehensive decision-making framework empowered by AI is constructed according to the logic of “decision-making mode selection–decision-making process–decision-making outcome”, fully considering the overlap and complementarity of substitutive decision-making and cooperative decision-making, and emphasizing the importance of human-machine collaboration. Finally, future research is discussed from three aspects: the antecedents and consequences of AI decision-making, the design of decision-making mechanisms, and the impact of China’s unique context, providing reference for promoting the transition of AI from low-level perceptual intelligence to high-level decision-making intelligence.

Key words: AI; substitutive decision-making; collaborative decision-making; human-machine collaboration

(责任编辑: 宋澄宇)