

AI 时代人机关系的演化和心智重塑*

程思琪¹, 喻国明²

(1. 暨南大学 新闻与传播学院, 广东 广州 510632; 2. 北京师范大学 新闻传播学院, 北京 100875)

[摘要] 人工智能技术的广泛应用正在深刻重塑人机互动的的基本模式。人机关系正从单一的工具性协作, 逐步演化为涵盖执行协作、认知协同与情感支持的多元互动形态。基于分布式认知理论, 不同的人机互动模式对人类认知能力具有双重影响: 一方面, AI 能够扩张并增强个体的认知功能; 另一方面, 它也可能导致某些认知功能的收缩与退化。未来关于人机关系的理论探讨与人机互动的实践设计, 均需探索这种张力的平衡与调节机制, 发挥人机关系对人类心智的促进作用, 抑制人机交互对人类心智的负面影响。

[关键词] AI; 人机互动; 分布式认知; 心智重塑

[中图分类号] B842 [文献标识码] A [文章编号] 1008-1763(2026)03-0150-09

Evolution of Human-Machine Relationships and the Cognitive Remodeling in the Era of AI

CHENG Siqu¹, YU Guoming²

(1. School of Journalism and Communication, Jinan University, Guangzhou 510632, China;
2. School of Journalism and Communication, Beijing Normal University, Beijing 100875, China)

Abstract: The widespread adoption of Artificial Intelligence (AI) is profoundly reshaping the fundamental patterns of human-machine interactions. These interactions are evolving from a single, tool-based collaboration model into diverse forms that encompass task execution, cognitive co-operation, and emotional support. From the perspective of distributed cognition theory, different modes of human-machine interaction exert a dual influence on human cognitive abilities. While AI can extend and enhance individual cognitive functions, it may also lead to contraction and decline in certain cognitive functions. In the future, theoretical discussions of human-machine relationships and practical designs of human-machine interaction should therefore explore how this tension can be balanced and regulated, so as to enhance the benefits of human-machine relations for the human cognition while reducing their negative effects.

Key words: AI; human-machine interaction; distributed cognition; cognitive remodeling

一 引 言

人工智能技术自 20 世纪 50 年代发展至今, 已从单一能力突破(如 AlphaGo 在围棋领域超越人类)演进至以 ChatGPT 为代表的通用整合与生成

阶段, 形成多领域、多层次、多场景应用的生成式人工智能新格局。随着模型推理能力的提升, 结合多模态信息处理与语音交互技术, AI 通过参与新知识生产, 逐渐获得类主体性地位, 对传统人类本位的智能观构成挑战。以新闻业为例: 写作机器人实现低成本高效的结构化报道生产, 算法分发提升内容精

* [收稿日期] 2025-05-07

[基金项目] 国家社会科学基金青年项目: 人与生成式人工智能交互下的信任度曲线与实践效应研究(24CWX041)

[作者简介] 程思琪(1992—), 女, 湖南岳阳人, 暨南大学新闻与传播学院讲师, 博士, 研究方向: 认知传播学、跨文化传播。

准度,多模态大模型重构新闻形态,人工智能自动化已成为重塑新闻生态的基础逻辑^[1]。大模型技术的可扩展性、多任务适应性与能力可塑性,使其广泛应用于对话交互、创意生成及知识管理场景。目前行业大模型已渗透金融、医疗、教育等领域,覆盖需求分析、模型选型、应用部署等关键环节。

当AI成为日常生活的行动者、组织者乃至决策者,人类语言、关系与认知边界正经历结构性变革^[2]。彭兰指出,智能体如同映照人类认知与社会生态的镜子,因此AI研究需超越人机差异比较,聚焦智能体对人类的重塑机制^[3]。随着大模型构建起拟像世界的体系,AI的类人属性日益丰富,促使人类形成涵盖技术信任、准人际信任等多维度的复合认知结构^[4]。人机情感互动的常态化更催生出兼具亲密性与异质性的“超社交关系”^[5],建立起跨物种、跨物质形态的新型交往范式^[6]。至此,人机互动已成为继人际关系之后的基础社会关系。例如“人机双螺旋”内容创作模型提出,人类与AI通过双向增强实现协同进化,其流程涵盖共情、提示、构思、原型、测试与执行六个阶段^[7]。具体表现为:人类借助数字角色分析用户需求,通过提示工程激发AI创意生成,利用大模型快速可视化原型,再基于AI模拟能力优化迭代。这种深度协作标志着人机关系从工具性使用转向共生性参与。

分布式认知理论提出,认知不仅仅发生在个体的大脑中,而是分布在个体、外部工具、符号系统和社会互动之间,也即认知不仅限于个体思维,还涉及与他人思维、物理对象和技术的互动^[8]。当人与AI互动时,思考和决策的过程延伸到AI的算力、算法模型和界面设计中,整个认知过程通过人类自然语言表达—AI大语言模型处理—交互界面输出三个层级的信息流动来完成。人机交互研究本质上仍是对交互情境中人类心智的考察,需从情感、认知、行为三维度观测其新变。本文将从人机互动模式的演变与新型人机关系的分类着手,基于分布式认知理论来探讨AI时代人机关系对人类心智能力的重塑机制,最终提出AI时代人机关系的学术思考路径与实践应用启发。

二 AI时代的人机互动模式 与新型人机关系

(一) 人机互动模式的演变

早期对人类机器关系的现象学考察将“机器”视作为人在日常生活中接触的各类物理器械,包括简单

工具(粉笔、闹钟等)与复杂机器(天文望远镜、飞机等),人机关系即人在借助各种机器体验世界时所产生的具身关系,是机器作为人类身体延伸去体验世界的关系^[9]。在这种人机关系的概念下,电视、电影、广播等媒介都可以被视为受众体验世界的工具,通过媒介体验的世界(媒介景观)与直接体验的世界(现实景观)共同作用于人们对外界和自我的认知。和其他工具不同的是,人与媒介的关系除了这种媒介充当中介的具身化体验,还有媒介扮演拟主体性角色的准社会互动体验,即媒介被当作可以进行社会交往的“他者”。例如,观众会对电视剧中的角色产生互动错觉与单向的情感依赖,这种关系被称为准社会关系^[10]。计算机应用的互动性进一步深化了人机关系中的社交属性,媒介等同理论认为计算机用户会将表现出一定拟人线索的计算机看作社交对象,并复现人际传播中的社交规则^[11]。从上述关系演进路径来看,随着非人类实体的拟人化程度增加,人机关系与人际关系之间的差异越发模糊,这种模糊性在人工智能应用普及后更加凸显。

自动化机器时代的劳动分工是将常规数据收集和分析任务分配给AI,将问题提出与结果解释等高阶任务交由人类负责,但现在人工智能大模型已经可以承担更广泛、更深度和更需创造力的高阶任务^[12]。传统人机交互的重点放在人类和非AI计算系统之间的相互作用上,这些系统依赖于固定的规则逻辑和算法来响应来自人类的输入,人类以“刺激—反应”的形式与系统交互,系统在人类监控下执行指定任务,本质还是承担一个工具性角色。在出现AI自主性驱动的人机系统后,机器从辅助工具拓展为与人类协作的队友。例如,大模型凭借理解复杂语言模式和产生类似人类反应的能力,已经可以与作为提示者的人类共同完成知识的创造和传播。

当AI被视为协作伙伴时,人们会期望像人际沟通一样从协作者处获得回应和反馈,尤其是情感反馈。人工共情,即AI感知、理解并共享他人思想和感受的能力,被认为是弥合人类在情感和社交层面的客户体验差距的关键,应该成为未来AI驱动的企业与客户互动中的重要考量因素^[13],例如人工智能在营销领域就存在巨大机遇和潜在收益。缺乏共情能力的AI往往很难用于涉及直觉和感情的任务,难以通过计算模型实现对人类认知和情感共情的编码,AI代理可以通过展示共情的表征,包括观点采择、共情关怀和情绪传染,来模拟人类共情。在

适当的情景中对用户实施人工共情,可以更好地满足用户的情感需求、提高关系满意度、提升用户幸福感,这对市场营销活动尤其有益,可以优化客户体验。当前 AI 在营销中的应用大多集中在影响客户体验的认知和行为维度上,例如使用算法推荐产品以减少信息过载,但是暂时忽略了跟客户的情感和社交互动,存在回应的机械化与沟通的低效性等问题,这恰好是 AI 营销与人类营销的差距所在。如果 AI 在人机互动中可以增强积极情绪或调节消极情绪,就能够更好地创造人机沟通中良好的情感循环。一个具有共情能力的 AI 代理可以通过观点采择来理解情景,并利用共情关怀来识别消费者的情感状态,为其提供情感支持,这些行动可以显著提高消费者黏性^[14]。

当前备受关注的**人机互动领域**主要包括工业生产、医疗健康、商业运营、教育实践、环境保护及创意工作等。其中,技术部门通过人工智能与机器人研究,持续优化价值链的效能、动态性、灵活度与适应力;在决策科学领域, AI 系统被用于评估高危情境、提供分析建议,整合现有战略技术成果,实现知识共享与资源利用最大化;在沟通交流层面,社交机器人帮助孤独症儿童提升社交技能,算法推荐与 AI 客服在接待服务与旅行服务领域提高了工作效率,教育助手通过个性化指导提升学生学习效率;对于设计师、科研人员等创意思维工作者而言, AI 在构思理念、生成执行方案等方面展现出独特价值^[15-17]。由此,从认知决策到情感交流,人机交互开始在越来越多的行业日常化。AI 技术的发展趋势有两个典型方向:一是技术能力从专用智能走向通用智能,二是其自主决策能力由低变高。越先进的 AI,越能在无持续人类干预下,通过感知环境、分析信息来制定决策并执行行动。根据 AI 技术能力与自主决策能力的发展程度,当前人机协作活动中 AI 已经可以在行为、认知和情感三个层面承担不同的角色分工。据此,基于功能导向可以将人机关系划分为执行协作关系、认知协同关系与情感支持关系。

(二)新型人机关系的分类

执行协作关系是人机交互初级阶段出现的关系类型,根据自主决策能力的高低,执行协作关系中的 AI 可以分为工具型 AI 和自主型 AI。决策能力低的工具型 AI 表现为人类发布指令, AI 理解人类意图并作出符合社会规范的价值响应,大致包括:信息处理型工具,如搜索引擎与智能翻译;生产力增强型工具,如编程辅助与办公自动化;决策支持型工具,

如金融分析与医疗诊断。它已广泛应用于医疗健康、工业生产、家庭服务等领域,如为老年人与残障人士提供辅助的护理机器人、制造业中的协作机械臂、教育场景的学习机器人、金融行业的数据分析工具等^[18]。决策能力高的自主型 AI 能在动态环境中独立完成价值判断与行动选择,其具备感知独立性、决策闭环性与适应性学习等要素。自动驾驶是典型的自主型 AI 应用,计算密集型 AI 技术的发展使自动驾驶车辆实现了高精度环境感知,能在无人工干预下作出安全的实时决策^[19]。随着 AI 环境感知、自主决策、执行任务能力的不断增强,其自主决策应用从自动驾驶拓展到无人机、智能农机、物流机器人等多个领域。例如,无人驾驶收割机越来越具备自动收割并优化农田管理策略的能力,无人战斗机可以自主侦察并执行战术决策。与工具型 AI 相比,自主型 AI 更具备通用智能的表征,但本质上依然是专用智能的集群化扩展。

大语言模型让 AI 表达的流畅度达到前所未有的水平,使得人与 AI 的认知协同关系得以产生。从现实应用来看,当前最具代表性的协作型关系是在辅助创意写作方面人与 AI 成为共同作者。ChatGPT 等大语言模型应用凭借其语言能力(生成流畅文本)与构思能力(产生新创意)与人类协作者协同创作。换言之,当前认知协同型人机关系正从会议安排、拼写检查、航班预订等简单机械任务应用场景,逐步向复杂对话、生成文本、创作高质量图像视频,以及为网站和应用制作交互式 UI 原型等任务拓展,极大地提升了诸多行业的工作效率。整体而言,人类—人工智能的认知协同主要从三个方面支持知识生产:日常任务自动化、增强认知能力与促进创新。具体而言,大语言模型可以通过简化各种研究任务、放大分析能力,帮助人类产生新的见解和想法^[12]。基于理论—实践与 AI—人类两个维度,当前人机的认知协同表现出四种类型:一是 AI 主导理论构建(AI-dominant theorizing)(理论性高, AI 参与度高),大模型在理论研究中起主导作用,以最少的人类干预驱动理论见解的生成和分析;二是 AI 作为实践者(AI-artisanship)(实践性高, AI 参与度高),在人类指导下执行实际任务,应用于现实场景;三是 AI—人类共生(AI-human symbiosis)(实践性高,人类参与度高),人类与 AI 协作互补,人类创造力与 AI 分析能力协同作用;四是人类主导 AI (human-dominant AI)(理论性高,人类参与度高),人类在 AI 辅助下主导概念建构和理论发展^[12]。

情感支持关系是人机深度交互后出现的关系类型。情感曾被认为是人类区别于AI最根本的特质,但伴侣型关系的出现意味着人机交互从认知沟通衍生到情感交流。从需求端来看,人类在追求与AI无障碍互动的过程中期待与AI建立类似于人际关系的亲密关系;从供应端来看,AI技术已经可以模仿人类情感。在供需的合力作用下,人机情感互动逐渐趋向日常化^[20]。这种建立在复杂技术系统基础上的情感互动超越了类社会互动,更接近于人际社交关系模式,但又不完全等同于面对面的人际交往。有学者打破“拟人化”思维局限,将其解释为跨生命、多物种的新交往模式^[7]。例如,通过让聊天AI充当朋友、恋人等多种角色,许多用户与AI建立了比与人类更深厚的情感联系,并产生了强烈的情感依恋。随着日常互动的增多,人类可能会在交流中受到聊天AI的影响而产生一些情绪变化。如果聊天AI的表达中出现不道德、含有威胁、侵犯隐私或其他伤害性行为,会导致用户出现恐惧或悲伤等负面情绪,长此以往可能会影响其心理健康。例如,在使用聊天机器人Replika的过程中,Replika会突然输出一些过分亲密的语言,让用户觉得突兀和反感^[21]。当然,如果聊天AI能够提供快乐和友好的交流,人们也会获得相应的情绪支持。

但是,执行写作、认知协同与情感支持三种人机关系的分类不具有绝对性,而是一种依赖于人类意图、使用场景和交互深度的相对性产物,三者的边界具有流动性、模糊性和交织性。三种关系可以被想象为一个三维空间模型,其中X轴即执行协作关系,表示从低到高的任务完成度;Y轴即认知协同关系,代表从低到高的思维耦合度;Z轴即情感支持关系,表示从低到高的情感联结度。在这个空间中,任何一段具体的人机关系都可以被映射为一个坐标点,大多数具有意义的人机交互状态往往位于空间内部,是三个维度不同程度叠加的区域,而非单一轴线。基于这一空间化的构想,我们可以通过具体应用场景的坐标定位来进一步理解三种关系如何在现实的人机互动中相互交织。例如,护理机器人可以被定位在“高执行中认知高情感”上,因为它既能够完成具体操作,又能提供信息解释,同时给予情感陪伴;编程助手则更接近“高执行高认知低情感”的位置,强调代码生成与逻辑协同,但在情感维度上相对缺乏;而AI聊天伴侣则体现为“低执行中认知高情感”,其任务执行较弱,但是可以在对话交流和情感支持上发挥重要作用。由此说明,执行协作、认知协

同与情感支持三类人机关系并非互斥的类别,而是在不同情境中以不同组合被激活,呈现出并存交织的关系状态。

三 分布式认知理论下新人机关系 重塑人类心智的可能性

(一) 分布式认知理论

分布式认知理论认为认知不仅存在于大脑,更存在于世界^[22],认知活动并非仅由单个个体完成,而是在一个由人、工具、环境组成的协作系统中协作完成,功能认知系统的四个单元(感知、记忆、处理和执行)分别由不同行动者分别承担^[23]。分布式认知理论作为研究人类群体协作的框架,通过将认知疆域从个体内部拓展至人与环境资源的交互层面,为理解新人机关系对人类认知的影响提供了可行思路。具体而言,认知活动不再局限于个体大脑中,而是在人与AI系统交互之中协同完成。AI与人共同承担信息加工、判断和问题解决的任务,形成一个混合智能体系统。当认知系统以社会分布式形态存在于多个相互依存的行动者——如人与AI——之间时,个体间需要通过协作整合各自的问题解决资源,这些资源的动态协调机制对其协同行动至关重要^[8]。人类协作能力的核心要素是心智理论(Theory of Mind, ToM),即智能体对自身及他人心理状态推断的能力,传统的关注重点是尽可能使AI具备人类水平的ToM推理能力,但是这类目标导向难以识别人类在协作任务中优于AI的关键要素,并且忽略了协作环境对人类认知实践的重要意义。分布式认知理论将研究重点从“以人为中心”转向去中心化的“关系本体论”,将AI视为认知系统的核心参与者。

在这一前提下,分布式认知理论提出了两个基础问题:通常归属于个体心智的认知过程如何在群体中实现?参与群体活动如何反塑个体心智的认知特性^[24]?人类心智能力通常包括情感、认知和行为三个层次,其中情感最易受影响,认知相对稳定但可塑,行为作为心智系统最外显的层次变动最慢。但是AI模仿人类心智的顺序呈现出一个与上述路径的逆向结构:行为(走路、开车等)作为可编程、目标导向的物理或数字行为,最容易被模拟;认知模拟诸如知识整合、推理创作等,虽然不等于人类认知过程,但可以有效呈现产出结果;情感(共情)因为与自身经验、生理基础密切相关最难模拟,AI可生成“共

情话术”，却难以产生真实情感状态。AI时代，人类心智能力的三个层面都能在一定程度上通过执行协作、认知协同与情感支持三类关系与AI交互，在交互过程中个体心智被影响甚至重塑，总体而言，处于人机关系中的个体表现为心智能力扩张与收缩两个面向。

(二) 心智能力如何被重塑

扩展心智理论(Extended Mind Thesis)提出，若外部工具(如笔记本)在功能上与内部认知过程(记忆)等效，在使用上也具有稳定性、即时性和可靠性，那么该工具应被视为认知系统的组成部分。心智作为一个遍布身体与环境的动态系统，其中具身性、行动性与环境适应性共同构成了认知的本质^[25]。Smart将扩展心智理论应用于当代数字技术与互联网语境，指出网络与AI等数字媒介作为外部认知基础设施参与人类推理、记忆与决策，重塑心智功能的边界^[26]。当外部技术被持续地、可靠地嵌入人的认知链条时，它开始真正承担某些认知和情感功能，进而重塑人类的心智能力。已经有认知神经研究揭示了神经可塑性，长期使用外部工具的个体会重塑大脑的认知加工路径，例如使用搜索引擎后，人更倾向于记住信息的位置，而非信息本身，这对应海马体与前额叶的活动变化^[27]。

当AI认知决策参与的范围越来越广，人机情感联结日益深入，人机共生的关系网络逐渐替代部分传统的人际关系网络。在这种背景下，AI不仅可以影响人类在成长期的社会化过程，也可以改变人类成长期过后的心智能力。这种转换可能会产生镜像效应，即人类反向参考AI的决策机制，一方面可能由此打破自身的认知局限，另一方面可能因此出现围绕AI数据驱动逻辑的认知同质化危机与人类认知权威解构，人类心智能力既可能被扩张，也存在被收缩的风险。

心智能力扩张指的是人通过与AI的互动，在信息处理、思维广度和情感调解等方面获得了超越个体本身的能力，相当于一种心智外延或者认知带宽的增容。AI可以作为心智“外接器官”，通过提升心智的灵活性、广度和韧性，帮助人类跨越认知与情感的局限，增强复杂任务处理、自我调节等能力。例如，在执行协作关系中，AI分担重复性任务可以将人类从机械劳动中解放出来，投入更复杂的创造性活动；在认知协同关系中，人类通过与AI的对话突破认知盲点，形成更开放的解释框架；在情感支持关系中，AI通过缓解人类的焦虑、孤独与轻减认知负

荷，增强人类的决策能力。

心智能力收缩是指人类在过度依赖人工智能或外部技术时，部分认知功能不再得到充分调用或训练，导致心智能力在范围或深度上出现弱化的过程。麦克卢汉在《理解媒介：论人的延伸》中曾有类似论述：新技术延伸某一感官时，往往会使其他感官麻痹，例如记忆外包给文字后，个体记忆力会消退^[28]²⁵⁻²⁶。如果将“延伸”视为心智能力扩张的哲学表述，“麻痹”则呼应了收缩的深层逻辑。在人机关系中，当执行任务被AI替代时，人类技能逐渐失去练习与调用的机会；当推理和记忆被外包时，批判性思维与独立判断趋于弱化；当情感支持依赖于拟人化机器时，真实的人际交往动机可能被侵蚀。

心智能力的扩张与收缩是人机交互中并存的孪生现象，看似矛盾，实则构成一种内在的双向动态。在元认知层面，即个体对自身认知活动的察觉、监控与调节过程中^[29]，AI一方面通过提供即时反馈与外部参照，帮助人更清晰地意识到自身局限，从而强化自我监控与学习调节；另一方面，当个体习惯于无条件依赖AI时，其对思维过程的检验与修正被外包，批判性监控能力因此弱化。事实上，这种扩张与收缩的张力，已在执行协作、认知协同与情感支持等多种人机交互场景中反复显现。

四 新人机关系下人类心智能力扩张与收缩的具体表征

(一) 心智能力的扩张：行为偏差纠正、认知边界拓展与情感投射对象延展

心智能力的扩张并非单一维度的增强，而是沿着执行、认知与情感三条路径展开，对人类行为模式、知识边界与社会关系产生深刻影响。

在执行协作层面，AI对人类心智的纠偏效能本质上源于其精准的功能聚焦，这种优势来自AI与人类在变量处理能力和决策稳定性上的差异。人类的工作记忆仅能处理约七个信息单元，面对复杂决策时容易过度强调显性变量，忽略隐性因素，而AI的高信息容量更容易实现全维度决策。并且人类的噪声过滤能力相对较弱，情绪波动、时间压力、社会期望等干扰变量都会影响决策稳定性。AI将人类从这种生物学限制中解放出来，减少人类认知与行为偏差，优化决策效率。另外，这种关系也代表着人类在部分任务场景下决策权的让渡，推动人类认知从“理解世界”向“理解智能体如何理解世界”转换。

认知协同更突出地表现为 AI 对心智能力边界的拓宽,因为人类可以通过 AI 实时调取远超脑容量的知识库。这种关系的本质是 AI 通过机器学习不断逼近人类认知的“暗知识”(tacit knowledge)边界,人类则通过价值判断和创造性思维引导 AI 的发展方向。当前讨论热度较高的主题是接受高等教育的学生群体如何在学业中使用 AI,尤其是 AI 如何影响写作这一复杂认知行为。AI 最初的作用主要体现在利用语法校对、风格检查等提升写作质量,随着自然语言处理技术的突破,新一代大语言模型已将 AI 的学术写作辅助能力拓展至选题构思、框架搭建、文稿撰写及修订润色等范畴,能够为写作全流程提供多维度支持,提升写作效率与表达清晰度。对于学术写作中至关重要的批判性分析与反驳论证等环节,AIGC 已经展现出激发创作灵感、增强文本连贯性、提升整体质量的潜力^[30]。研究发现,与 AIGC 进行高度互动、迭代式协作的博士生能取得更优异的协作表现,这证明有必要在高等教育中深化人机认知协同^[31]。

AI 在情感支持上的作用不仅是技术的整合,更是对社会关系认知的范式变革。人们将 AI 视为能够进行情感互动的关系性实体,将对同类的情感投射延展至 AI。尤其随着 AI 可以提供跨越时间的延续性情感交流与体验,技术模拟与真实情感体验的边界逐渐消弭。当前人机亲密关系主要有两种方向:一是 AI 作为数字纪念碑,用于保存、延续、模拟逝者的记忆、情感与互动,成为“后人类”亲属,重塑人类对死亡的认知^[32]^[25];二是 AI 被用来填补亲密关系缺口,尤其在人口老龄化和家庭结构碎片化、存在较大照护缺口的当下,AI 作为护理角色,通过日常陪伴、共情倾听等情感劳动,建立起不亚于人际关系连接强度的互动关系。AI 越能模拟人类信念、情感和能动性,就越能影响人类的身份认同与自我认知。

(二)心智能力的收缩:认知控制的外包,判断力的削弱与情感联结的弱化

与心智能力扩张相对应的是人类在认知控制外包、判断力削弱与情感联结弱化三个维度上心智能力的收缩。

首先,在执行协作的场景中,自动驾驶最能体现认知控制的外包。司机长期依赖车辆的自动决策系统,会削弱对路况的主动监控与应急反应能力,本质上是放弃了对整体认知过程的主导权^[33]。这种外包由两个机制共同推动:其一是认知卸载,即个体通

过外部工具降低认知负担,以追求效率。大语言模型在知识生成上的卓越性能使其在写作、问题界定等任务中被广泛使用,但这种转移也带来去技能化风险。研究发现,当工作者从主动参与角色转变为被动监督角色时,可能会出现技能流失现象^[24]。这引发了对人机协作中基础知识与思维习惯是否会因过度卸载而退化的担忧。在人机交互时,人类用户应当进行多大程度的认知卸载,以及让渡哪些认知功能,尤其是决策功能,都需要仔细考量。其二是新增的认知负荷。自动化反讽论揭示,人机合作常伴随悖论:一方面,自动化系统的设计缺陷本身可能成为故障源;另一方面,未被完全自动化的任务仍需人类承担,造成责任不清。在 AI 辅助设计中,设计师常需投入大量精力验证和修正生成内容,这反而延长了迭代周期。同时,AI 目前还存在简单任务增效显著、复杂任务效果受限的缺陷,例如,可以处理简单的法律分析,却在复杂法律推理中频频出错^[34]。这种由技术复杂性带来的额外判断与理解成本,迫使用户因负担规避而更加依赖 AI 的结论,从而加速了认知控制的外包。

其次,在认知协同与决策支持领域,AI 的广泛应用则导致判断力的削弱。数据驱动的算法已进入社会福利、司法、医疗和教育等关键公共领域,甚至在某些环节部分取代人类判断。然而,人类在面对算法建议时往往表现出自动化偏见(明知 AI 有风险却依然选择听从)以及选择性服从(当算法结论符合既有刻板印象时更愿意采纳)^[35]。这意味着,如果缺乏批判性监督,算法治理可能放大偏见与错误。尽管“以人为中心的设计”(human-centered design, HCD)理念通过责任性(确保 AI 决策可追溯)、自主性(动态环境中的自主感知与决策)、公平性(确保模型不因种族、性别或残疾等因素产生歧视)、透明度(向终端用户公开模型运作机制)、可解释性(阐明模型决策的内在逻辑)、可信度(提供安全保障)与隐私保护(防止数据未经授权使用)等原则,提高了 AI 辅助决策的质量^[36],但这种改进也增强了人类对算法结果的依赖。在提升效率与信任感的同时,人的独立判断力可能因此进一步被削弱。

此外,同化效应认为,个体在智能体身上感知到的“人性”越多,在真实他人身上的“人性”感知就越少,进而弱化与现实社会的情感联结。AI 技术不断突破表层拟人化,力图赋予其类人情感属性,使其能够通过面部表情、肢体语言与语音语调等非语言线索识别并适应人类情绪^[37]。伴随类人化交互应用

的普及,情感因素在人机交流中发挥了增强沟通、增进信任和提升满意度的作用,有研究发现,人们在AI心理咨询师面前的自我揭露程度与面对人类咨询师相当^[38]。然而,其他研究表明,AI会改变人类的沟通方式,使人们在与真实他人交流时不自觉地沿用与AI对话的简化习惯,导致表达风格与情感语言趋于同质化^[39]。与此同时,AI的碎片化、顺从式回应提供的只是浅层情感体验,这可能削弱人们处理复杂人际情感的耐心,进而降低对冲突的容忍度,不利于长期关系的维系,并放大社交恐惧。例如,儿童在与语音助手交互中使用命令式语气时,这种攻击性表达会无意识地渗透到与亲友的交流中^[40]。

尽管智能体被赋予类人情感,但其情感水平仍远低于人类。然而,当其被感知为接近人类情感能力时,反而可能降低用户对他人“人性”的评价,诱发“去人性化”现象^[41]。这种对他人理性意图和情感体验能力的否定,会影响个体的社会行为模式,表现为冷漠、苛责、减少助人为乐行为,甚至增加暴力或不道德行为^[42]。例如,与聊天机器人的互动会降低人们的慈善捐赠意愿,并促使其保持社会距离^[43]。社会认知理论与心智理论进一步解释了这种AI对人机互动产生的溢出效应:社交型AI表现出来的社会属性可以成为用户学习社会行为的示范模型,参与社会规范的塑造^[44]。虽然AI在老年人、单身群体或儿童群体中填补了情感需求缺口,但这种依赖技术弥补社交陪伴的模式本质上反映了社会亲密关系的缺位。一旦人机亲密关系在更大范围内扩展,它必然会反过来重塑人际互动的准则。

(三)心智重塑的情境调节性:技术、文化与社会脉络的差异化影响

心智能力在不同人机关系中的扩张与收缩机制具有普遍性,但是这一过程不是在真空中发生的,而是嵌入技术成熟度、文化价值观与社会制度等情境之中,展现出显著的情境调节性。AI技术处于引入阶段时,更多地通过提高信息处理效率带来心智能力的扩张,但是随着其融入日常生活,认知控制外包、认知技能退化等风险逐渐显现。已经有学者发现,可解释性AI的设计与评价普遍依赖西方个体主义社会的需求假设,对集体主义文化背景下用户的偏好与使用逻辑缺乏充分关注^[45]。类似的偏向也出现在关于“算法厌恶”的研究中,目前的证据几乎完全基于西方社会,难以判断这种现象是否具有跨文化普遍性^[46]。这意味着人机关系对人类心智

能力的影响也可能并不完全遵循单一规律,而是受到文化价值观与社会规范的调节。这种差异也会折射到社会制度层面,例如福利导向型社会倾向利用护理机器人弥补照护缺口,人机关系在情感维度上的作用权重更大;相对而言,市场驱动型社会更注重生产效率,心智能力的重塑更集中在执行与认知层面。未来研究需要在跨文化比较中揭示这些差异,以构建出更具情境敏感性的心智重塑理论框架。

五 结论与展望:理论拓展与实践指引

本文基于分布式认知理论,探讨了AI时代的新型人机关系对人类心智的重塑机制:一方面,AI在执行协作、认知协同与情感支持三个维度上不断扩张人类的心智能力,帮助个体克服生理与认知局限,实现更高效的任务完成、更广阔的知识边界与更多样化的情感体验;另一方面,这种重塑也伴随着心智能力收缩的风险,包括认知控制的外包、独立判断力的削弱以及人际情感联结的弱化。扩张与收缩的并存说明人机关系对人类心智的影响是一种动态平衡展开的过程,也揭示了人机关系的内在张力。因此,未来关于人机关系的理论探讨与人机交互的实践设计,均需要探索这种张力的平衡与调节机制,发挥人机关系对人类心智的促进作用,抑制人机交互对人类心智的负面影响。

(一)拓展人机关系研究的理论框架

本文通过揭示人机关系对心智能力的双重作用推动了分布式认知理论的深化:它不再只是描述心智如何分布于个体—工具—社会网络之间,而是进一步强调了动态双向性与张力。基于此,本文提出三个未来研究路径。

首先是AI的技术框架效应研究。框架理论指出,新闻信息的组织、呈现和解释方式会影响受众对议题的理解和感知,同一信息在不同框架下呈现时,受众会产生不同的理解和决策^[47]²¹。当算法成为新闻信息的组织者、呈现者和分发者时,有必要考虑算法新闻是否同样存在框架效应,出现框架生产自动化,放大人类新闻从业者不自觉的偏见框架。因此,人们心智加工的对象是否已经是算法框架效应后的结果,是新闻传播学科需要重点关注的问题之一。此外,许多看似由人类产出的信息,实际上是由AI生成的。在这种情况下,AI的技术逻辑可能对社会各个行业人们的信息加工产生隐性框架效应,从而重塑人类的思维与决策方式。

其次是人机信任的动态性研究,以及AI拟人化与主体性研究。当AI对人类智能的还原度越来越高、人与AI关系日益紧密,人机信任关系的复杂性和重要性都在递增。AI信任研究既能指导可信AI的开发,也能促进人机交互与协作的良性发展,降低误用和弃用的风险。具体而言,可以基于社会认知视角来验证并提升人际信任的理论框架在人机信任中的适用性,也可以依托包含信任方、被信任方及交互情景的三维信任框架,验证影响AI信任的关键因素。拟人化研究则包含两个维度:一是人们如何感知智能体的拟人化程度;二是当人们认为AI在某种层面已经高度拟人化后,如何看待AI的主体性。结合当前AI拟人化发展程度与人们对AI主体性的感知,人机亲密关系似乎可以替代人际亲密关系,这正在慢慢颠覆人们对AI主体性的判断与看法。尤其具身AI进一步通过物理存在、动作自主性、情感表达丰富性更容易触发人类的类心智模型结构,对AI产生“主体幻觉”,进而产生亲密度更高的情感依赖。

(二)AI设计规范与AI素养提升的启示

人机关系对人类心智能力的重塑不仅是理论层面的问题,也深刻影响着AI技术在社会中的实践走向。只有在技术设计与用户素养两个层面同时发力,才能真正实现“趋利避害”,发挥AI的积极效应。在设计层面,AI不应被单纯理解为效率工具。当前大多数生成式AI系统仍然存在明显的“黑箱效应”,其决策过程缺乏可追溯性与可解释性,削弱了用户对AI建议批判性检视的能力,也可能在无意识中加剧对系统的过度依赖,放大心智能力收缩的风险。因此,未来AI设计应该兼顾透明性与可解释性,将模型的运作逻辑、数据来源等信息以可理解的方式反馈给用户,保障用户的认知自主性。在用户素养层面,AI的普及意味着个体必须充分意识到日常使用对其认知与情感的双重影响——既能提供能力扩张,也可能诱发过度依赖与独立思考的弱化。因此AI相关的培训与教育不应局限于技能操作,也需要展开素养导向型的教育实践,这不仅包括帮助用户理解AI的工作机制及潜在局限,还应引导其在不同任务中合理分配认知控制与外包比例,同时提升其在伦理判断与责任承担方面的敏感度以实现AI的理性使用。

[参 考 文 献]

[1] 常江. 人工智能时代的新闻行动者:人机比较与未来生态[J].

新闻与写作,2024(10):36-45.

- [2] 全燕. 人机传播的概念性边界研究:沿算法传播的分析思路[J]. 新闻界,2024(9):15-26.
- [3] 彭兰.“镜子”与“他者”:智能机器与人类关系之考辨[J]. 新闻大学,2024(3):18-32+117-118.
- [4] 陈慧敏,朱姿伊,金兼斌. 知识生产视域下生成式人工智能信任的复合框架、作用机制与未来进路[J]. 中国编辑,2024(12):60-69.
- [5] 牟怡,吴宇恒. 从准社会关系到人机关系:基于原真性与互动性的双维度归类模型[J]. 当代传播,2024(3):9-14.
- [6] 宋美杰. 走出人类世:作为同伴物种的人工智能与跨物种交往[J]. 新闻界,2024(7):14-25.
- [7] 周慎,莫菲菲,张佳伟. 人机耦合体与人机双螺旋:AIGC时代数字内容创作关系与方法重塑[J]. 数字出版研究,2024(4):121-129.
- [8] Perry A. AI will never convey the essence of human empathy[J]. Nature Human Behaviour,2023(11):1808-1809.
- [9] Ihde D. The experience of technology; human-machine relations[J]. Cultural Hermeneutics, 1975(3):267-279.
- [10] Rubin A M, Perse E M, Powell R A. Loneliness, parasocial interaction, and local television news viewing [J]. Human Communication Research, 1985(2):155-180.
- [11] Reeves B, Nass C. The media equation: how people treat computers, television, and new media like real people and places [M]. Cambridge: Cambridge University Press, 1996.
- [12] Clegg S, Sarkar S. Artificial intelligence and management education: a conceptualization of human-machine interaction[J]. The International Journal of Management Education, 2024(3): 101007.
- [13] Liu-Thompkins Y, Okazaki S, Li H R. Artificial empathy in marketing interactions: bridging the human-AI gap in affective and social customer experience[J]. Journal of the Academy of Marketing Science, 2022(6): 1198-1218.
- [14] Prentice C. Enhancing the tourist experience with emotional intelligence[J]. Tourism Review, 2020(5): 733-744.
- [15] Obrenovic B, Gu X, Wang G Y, et al. Generative AI and human-robot interaction: implications and future agenda for business, society and ethics[J]. AI & Society, 2024(2): 677-690.
- [16] Peca A, Coeckelbergh M, Simut R, et al. Robot enhanced therapy for children with autism disorders: measuring ethical acceptability[J]. IEEE Technology and Society Magazine, 2016(2): 54-66.
- [17] Celik I. Towards Intelligent-TPACK: an empirical study on teachers' professional knowledge to ethically integrate artificial intelligence(AI)-based tools into education[J]. Computers in Human Behavior, 2023, 138: 107468.
- [18] Chiriatti M, Ganapini M, Panai E, et al. The case for human-AI interaction as system 0 thinking[J]. Nature Human Behaviour, 2024(10): 1829-1830.
- [19] Atakishiyev S, Salameh M, Goebel R. Incorporating explanations into human-machine interfaces for trust and situation awareness in autonomous vehicles[C] // 2024 IEEE Intelligent

- Vehicles Symposium: IV. Jeju Island: IEEE, 2024: 2948 — 2955. DOI:10.1109/IV55156.2024.10588812.
- [20] 苏涛,彭兰.生成式人工智能时代的传播与人类新境遇:2024年新媒体研究述评[J].国际新闻界,2025(1):53—70.
- [21] Li H,Zhang R W. Finding love in algorithms:deciphering the emotional contexts of close encounters with AI chatbots[J]. Journal of Computer-Mediated Communication, 2024 (5): zmae015.
- [22] Norman D A. Things that make us smart:defending human attributes in the age of the machine[M]. Reading: Addison-Wesley Publishing Company,1993.
- [23] Rogers Y,Ellis J. Distributed cognition:an alternative framework for analysing and explaining collaborative working[J]. Journal of Information Technology,1994,9:119—128.
- [24] Hollan J,Hutchins E,Kirsh D. Distributed cognition:toward a new foundation for human-computer interaction research[J]. ACM Transactions on Computer-Human Interaction,2000(2): 174—196.
- [25] Clark A,Chalmers D. The extended mind[J]. Analysis, 1998 (1):7—19.
- [26] Smart P. Extended cognition and the Internet[J]. Philosophy & Technology,2017(3):357—390.
- [27] Sparrow B, Liu J, Wegner D M. Google effects on memory: cognitive consequences of having information at our fingertips [J]. Science,2011,333:776—778.
- [28] 麦克卢汉.理解媒介:论人的延伸[M].何道宽,译.北京:商务印书馆,2000.
- [29] Veenman M V J, Van Hout-Wolters B H A M, Afflerbach P. Metacognition and learning: conceptual and methodological considerations [J]. Metacognition and Learning, 2006 (1): 3—14.
- [30] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[C]//Proceedings of the 38th international conference on machine learning. 2021:8748—8763.
- [31] Nguyen A, Hong Y, Dang B, et al. Human-AI collaboration patterns in AI-assisted academic writing[J]. Studies in Higher Education,2024(5):847—864.
- [32] Savin-Baden M. Postdigital humans: transitions, transformations and transcendence [M]. Cham: Springer International Publishing,2021.
- [33] Miller E K, Cohen J D. An integrative theory of prefrontal cortex function[J]. Annual Review of Neuroscience, 2001, 24: 167—202.
- [34] Obrenovic B, Asa A R, Oblakovic G. The use of ChatGPT in the workplace: a bibliometric analysis of integration and influence trends[J]. AI & Society,2026,41:655—668.
- [35] 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(1):153—169.
- [36] Nazar M, Alam M M, Yafi E, et al. A systematic review of human-computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques[J]. IEEE Access,2021,9:153316—153348.
- [37] Bakola L, Rizos N, Drigas A. ICTs for emotional and social skills development for children with ADHD and ASD co-existence[J]. International Journal of Emerging Technologies in Learning,2019(5):122—131.
- [38] Lucas G M, Gratch J, King A, et al. It’s only a computer: virtual humans increase willingness to disclose[J]. Computers in Human Behavior,2014,37:94—100.
- [39] Hohenstein J, Kizilcec R F, DiFranzo D, et al. Artificial intelligence in communication impacts language and social relationships[J]. Scientific Reports,2023,13:5487.
- [40] Garg R, Sengupta S. He is just like me: a study of the long-term use of smart speakers by parents and children[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies,2020(1):1—24.
- [41] Kim H Y, McGill A L. AI-induced dehumanization[J]. Journal of Consumer Psychology,2025(3),363—381.
- [42] Schroeder J, Kardas M, Epley N. The humanizing voice: speech reveals, and text conceals, a more thoughtful mind in the midst of disagreement[J]. Psychological Science, 2017 (12): 1745—1762.
- [43] 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(3):849—865.
- [44] Guingrich R E, Graziano M S A. Ascribing consciousness to artificial intelligence: human-AI interaction and its carry-over effects on human-human interaction [J]. Frontiers in Psychology,2024,15:1322781.
- [45] Peters U, Carman M. Cultural bias in explainable AI research: a systematic analysis[J]. Journal of Artificial Intelligence Research,2024,79:971—1000.
- [46] Liu N T Y, Kirshner S N, Lim E T K. Is algorithm aversion WEIRD? A cross-country comparison of individual differences and algorithm aversion[J]. Journal of Retailing and Consumer Services,2023,72:103259.
- [47] Goffman E. Frame analysis: an essay on the organization of experience[M]. Cambridge: Harvard University Press,1974.