

特色专题

从“智能互联”迈向“智能互协”

程子学^{1,2}, 王晨^{2,3}

摘要 随着大语言模型 (large language models, LLM) 的飞速发展, 多智能体系统 (smart synergetic networks, MAS) 的通信机制, 从以本体论 (Ontology) 为信息解释基础的传统协议 (如 KQML) 逐步转向以 LLM 驱动的协议体系。梳理了 Google 公司等提出的智能体互联协议 (agent-to-agent, A2A) 与 Anthropic 的模型上下文协议 (model context protocol, MCP), 以及两者结合形成的 MCP×A2A 协议栈, 指出其为智能体功能共享与工具调用提供了统一接口, 成为“智能互联网”的关键基础。然而, 该体系当前仍以功能互通为主, 难以支持多用户、多智能体参与下的深度共协与心理调节。为此, 提出基于人机共协计算 (human engaged computing, HEC) 的“智能互协网” (智协网) 架构, 首次在智能体协议中引入“身-知-心”3 层结构, 支持柔性、共感、多方可持续协作。系统可根据个体状态动态调整任务节奏与交互方式, 提升智能体对人类的理解与支持能力, 缓解传统系统在人机共协中的“理解差”“节奏冲突”等问题。通过教育辅导、远程医疗等案例, 展示智协网在用户体验、系统韧性与信任维护方面的潜力, 标志着智能体网络从“功能互通”迈向“共协交互”的转变。

关键词 人机交互; 人机共协计算; 多智能体系统; 智能体互联协议; 模型上下文协议; 智能互协网

1 AI 时代下的新型社会结构与智能互联网的必要性

随着大语言模型快速发展, 人工智能 (artificial intelligence, AI) 正驱动着以“智能化”为核心的新一轮工业革命的兴起。智能体 (agent) 作为人类与 AI 系统之间的关键连接器, 不仅承担信息传递与协作的桥梁角色, 更以其感知、理解与反馈能力, 成为推动各类场景实现智能化转型的核心力量。AI 正迈向多智能体协作与人机共感融合的新阶段^[1], 并展示了在资源受限条件下提升效率与推理能力的技术路径, 推动了其通用性与专用性的融合^[2]。这不仅催生了多种垂直领

域智能体的加速部署, 智能的范围也正被重视为融合感知、认知与情感的多样性系统^[3]。根据 OpenAI 公司的人工智能发展五阶段模型, AI 目前已达到了“对话”和“推理”, 正迈进智能体“代理”, 将逐步进入“创新”与“组织管理决策”的发展期^[4]。也因此, AI 的设计理念正转向以人为本、强调共感与共创架构的“与人共创主体”。

与这一趋势相呼应, 2025 年 4 月, 由 Google 公司牵头, 全球 50 余家企业与组织响应, 提出“智能体互联协议” (agent-to-agent, A2A)^[5], 与 2024 年 Anthropic 公司发布的“模型上下文协议” (model context protocol, MCP)^[6] 形成互补。这些协议为不同平台 AI 间的通信与协同操作提供了统一框架, 使得每个智能体不仅能明示自身功能, 还能主动调用其他智能体的能力以及外部工具, 实现任务分工协作的高度灵活

1. 会津大学计算机理工学部, 日本福島 9658580

2. 上海工程技术大学, 上海 201620

3. 高知工科大学, 日本高知 7828502

收稿日期: 2025-08-18; 修回日期: 2026-03-02

作者简介: 程子学, 教授, 研究方向为共协交互与智能体协议, 电子信箱: shigakutei@gmail.com

引用格式: 程子学, 王晨. 从“智能互联”迈向“智能互协”[J]. 科技导报, 2026, 44(7): 80-90; doi:10.3981/j.issn.1000-7857.2025.08.00056

化。这一背景表明 A2A 与 MCP 机制正催生出一个高度协同、可自组织的“智能体互联网络”(internet of agents)^[7], 智能体间可通过协议协作与资源共享, 具备感知、认知、决策与反馈能力, 广泛连接人、机、物, 实现从“信息互联”向“智能互联”的跃升。

然而, 在 AI 间协作机制日益进步的同时, 业界对于“人与 AI 之间的关系”的洞察与理论更新却相对滞后, 即人如何更为结构性地理解和参与 AI 的决策过程, 并与其实现共同成长的路径尚不明晰。例如, 当多个智能体根据人类模糊的指令同时展开行动时, 常常会引发社会资源的竞争、伦理冲突等问题。此外, 过度依赖 AI 所导致的“代理行为”泛滥, 亦可能削弱个体的主体性与社会协调能力。

针对上述挑战, 人机共协计算(human-engaged computing, HEC)提供了一个关键的理论框架。HEC 强调人与计算机系统(包括 AI 等)在任务执行中的协同分工与相互增益, 其目标不仅是提高效率, 更在于共协交互如何促进人类能力成长、认知深化与价值实现。换言之, HEC 是一个以“共进化”为核心的人机融合框架, 强调从“完成任务”转向“共同成长”。

为此, 本研究基于 HEC 的理论框架和 A2A 协议, 提出“智协网”(smart synergetic network)思想。在多主体协作、跨领域项目推进等复杂情境中, 依据各参与者的身体状态、知识技能、心理情绪等要素, 实现最优的合作机制、角色分担、资源配置与进度协调, 最终达到人机共协的目的。

通过该机制, 诸如参与者间的速度差异、沟通水平、认知落差、任务理解偏差等“不可见的摩擦”, 将由智协网中的多智能体动态协调与弥合, 为每一位参与者营造一个流畅、舒适、无摩擦的协作体验。这不仅是 AI 辅助的一种表现, 更是一种以“共情”为基础的人机关系再设计。进一步讲, 智协网以 A2A 协议为通信基础和智能互联层, 依托 HEC 理论中的一个重要架构“身-知-心”, 构建了“身-知-心”3 层智能体互协结构。各层之间的智能体群通过具备通信、协作与任务解决能力的协议机制和群内“心”的调整机制实现互协联动, 并支持跨层处理。此外, 智协网内部还内嵌面向协作的知识图谱(协作知识), 通过知识图谱与互协协议的结合, 实现对上述复杂人际关系动态与社会课题的深度响应与人机共协。

2 多智能体间互联的进展与挑战

2.1 智能体在人工智能发展中的角色

AI 正从单纯地回答提问, 通过推理解决问题的系统, 迈向具备综合感知与决策能力的智能体系统, 实现任务自主化与增强对不同环境适应性。其本质可理解为“智能助手或为智能代理人”, 在复杂任务中代表用户完成信息查询与决策执行。未来, 智能体将广泛应用于教育、医疗、产业等领域, 成为人机共生、共创系统中的关键技术。一般而言, 智能体包含以下 4 个核心模块。(1) 感知模块(perception): 接收环境输入, 如文本、图像、语音等。(2) 规划模块(planning): 基于输入信息进行逻辑判断和问题分析, 设定目标并制定实现路径。(3) 执行模块(execution): 将规划转化为行动, 调用外部资源和工具完成具体任务。(4) 记忆模块(memory): 存储历史经验, 供当前判断与未来优化使用。

这些模块并非各自孤立, 而是构成了一个动态联动的智能系统: 感知输入—推理解—制定计划—执行动作—反馈更新, 如此形成自我调节的闭环。

以海外旅行计划为例, 智能体可作为用户的“代理人”承担如下任务: 搜集目的地信息、签证要求、天气与交通状况(感知); 分析预算与时间偏好, 规划最佳行程(推理与规划); 代表用户执行订票、酒店预订、餐厅预约等操作(执行); 记录旅行偏好与历史经验, 用于优化未来行程(记忆)。通过这一系列活动, 智能体成为理解人、服务人、代表人行动的代理人。

2.2 回顾多智能体系统的演进与挑战

具备感知、决策、执行能力的各个独立智能体, 可构成多智能体系统(multi-agent system, MAS), 围绕目标, 实现任务分工与协作。例如, 在智能交通系统中, 每辆车可视为一个智能体, 它能与其他车辆智能体或交通信号灯智能体协同, 实现车流动态调度与路径优化, 提升整体交通效率与安全性。

20 世纪 90 年代, 知识查询与操作语言(knowledge query and manipulation language, KQML)的提出为 MAS 的通信与协调奠定了基础^[8]。KQML 为智能体间通信提供了语义单元和通信行为, 规定了“如何”交换命令、知识和目标, 而本体论则提供了“交换什么”的共同理解。此外, 通过运用共同本体, 实现了不

同智能体间知识的语义互操作性,并为后续 FIPA-ACL(agent communication language)等协议的发展提供了理论基础^[9]。

2002 年,佐藤等^[10]将心理学中的认知平衡理论引入 MAS,提出了一种用于解决学生与教师研究室之间配属匹配问题的说服机制。系统中,智能体代理用户根据其心理因素(如信任、兴趣、满意度)与其他代理展开交涉与说服,逐步调整态度,实现心理上可接受的匹配方案。该机制有效降低了传统人力调配的时间与成本,同时避免人为偏见,为实现更高效、人性化的匹配系统奠定了基础。

Ogino 等^[11]针对以云为中心的物联网(internet of things, IoT)架构所面临的通信负载、响应延迟和隐私泄露等问题,提出了一种引入多智能体技术的柔性边缘计算(flexible edge computing)架构。该方法在承担全局最优化任务的云端和负责实时控制的边缘侧,分别部署自治智能体,实现了云与边缘之间协同调节、动态平衡的“柔性网络(软网络)”。每个智能体可根据任务性质和资源状况灵活分配与迁移,从而调节局部最优与整体最优之间的权衡关系。

进入大语言模型时代后,以 LLM 为基础的智能体不仅能够设定目标、设计实现路径、进行分析与决策,还具备调用工具与反馈能力,从而使 AI 逐步从“对话”型演进为“代理”型。基于大型语言模型的多智能体系统(LLM-MAS)在社会行为模拟、心理互动仿真与个性化推荐等领域展现出显著潜力,每个智能体能够灵活再现多方对话与决策过程,在教育、政策制定等场景中具有良好适用性。然而,在团队协作、经济建模等动态性与一致性要求较高的应用中,目前仍缺乏系统性的性能评估基准。这类综合性基准的建立,将成为推动 LLM-MAS 走向现实复杂环境应用的重要关键^[12]。

为了适应多智能体间在复杂任务中的动态分工与交互需求,一系列开发和运行系统架构应运而生。(1) LangChain/LangGraph: 支持多智能体 workflow 构建,复杂状态管理和智能体间的串行、并行任务编排^[13-14]。提供基于图的灵活 workflow 设计和丰富的工具集功能,在动态智能体协作模式方面表现优异。(2) AutoGen^[15]与 CrewAI^[15]: AutoGen 提供对话驱动的多智能体会话框架,能够灵活地自定义和精细控制

功能,特别擅长代码生成和计算密集型任务。但复杂性较高,学习成本较大^[15-16]。CrewAI 是重视基于角色分工的自主 AI 智能体协作框架,重视基于角色分工的自主 AI 智能体协作。以结构化 workflow 和面向商业用户的易用性为特色,但在高度定制化方面灵活性有限^[15,17]。

群体智能系统的协作远不只是将多个智能体简单拼接起来。现实应用中常出现一系列协作障碍,例如,在用户服务、数据处理或系统控制等场景中,不同智能体往往基于不同的架构与语义模型运行。这就造成了接口不兼容、任务目标分歧、状态感知不同步等问题,进一步引发协作效率低下或系统行为不稳定。这种平台间“互不相识”的状况,使得资源共享与多角色合作面临重重障碍,也突显出跨平台智能体协作机制的迫切需求^[18]。上述挑战反映出,要构建高效、多角色协作的 MAS 系统,必须建立统一、可扩展的工具调用和通信与协调协议。这一背景促使 MCP 与 A2A 等新一代智能体调用工具和相互联结及功能共享的协议应需而生。

2.3 智能体协议 MCP 和 A2A 正成为智能体互联的基础

在多智能体系统的发展过程中,2 个关键协议正成为核心驱动力: 2024 年 11 月 Anthropic 公司推出的 MCP 与 2025 年 4 月 Google 公司主导的 A2A。前者关注智能体与外部资源的连接能力,后者则聚焦于智能体之间的通信互联能力。

1) MCP 与 A2A 的功能分工与互补。MCP 的使命在于使大语言模型跳出“孤岛运行”,通过结构化协议对接外部工具、数据库、知识库与各种应用程序编程接口(application programming interface, API)。它如同赋予 LLM“手和眼或比喻成智能 U 盘”,使其能在现实任务中主动获取所需信息并完成操作。其技术特性如下。(1) 高度兼容性: 与检索增强生成(retrieval-augmented generation, RAG)等主流机制良好协同。(2) 统一数据结构: 采用模式(Schema)规范请求与响应格式。(3) 上下文记忆能力: 支持模型对历史调用状态的记录与利用。(4) 安全控制机制: 包含访问权限、频率控制与数据加密。

与此同时, A2A 协议的设计目标则是打造“智能体之间的通用语言”。它通过轻量级、平台中立的 JavaScript 对象表示法(JavaScript Object Notation,

JSON)结构,定义函数调用(函数名+参数+返回值)和身份认证机制,使不同模型与平台下的智能体能够高效交互,完成任务转交与状态反馈。其应用范围日益广泛,成为了多智能体生态系统中不可或缺的“通信协议层”。

MCP 与 A2A 的结合构成了继互联网传输控制协议/因特网互联协议(transmission control protocol/internet protocol, TCP/IP)之后另一种类型的智能体互联网协议栈,将推动新一波跨不同 LLM 模型、跨任务的智能体协同应用浪潮,成为构建智能互联网的基础。然而,当前 MCP 与 A2A 的集成应用主要停留在“连接”与“调用”层面,没有达到深层协作层面。这一差距正是后文“智协网”构想的出发点。

2) MCP×A2A 智能体协议栈。2025 年 5 月, Jeong^[19]提出了一个更加系统化的 MCP×A2A 分层协同架构模型,以实现智能体与工具之间,以及智能体彼此之间的高效联动与动态协作。

其主要 4 层结构的组成与功能如下。第 1 层(最上层):智能体管理层。实现智能体的注册、功能编排、发现与调度;是协调系统运行的中枢,确保智能体之间合理分工与状态同步。第 2 层:核心通信协议层。在系统层面协调 A2A(通信)与 MCP(调用)的协同工作,负责定义任务调用结构、路由机制、处理反馈,并保证任务语义传递的一致性。第 3 层:工具整合层。通过 MCP 协议对接数据库、API、知识库、文件系统等外部工具资源,实现功能调用、数据处理与结构化反馈。第 4 层(最下层):安全认证层。提供身份验证、权限控制、日志追踪与端到端加密机制,确保系统运行的信任基础与安全。

用户可以通过协议栈连接多个智能体协同工作,各智能体基于 A2A 通信机制互相连接和互补功能、可以进一步通过 MCP 接入各种外部资源,包括数据库、API、工具链等,为智能体群提供执行能力。协议栈构建起“自然语言输入—协同任务规划—工具调用执行—多智能体反馈”的闭环。“MCP×A2A 分层协议栈”不仅清晰划分了各协议功能在系统中的职责,还为大规模智能体网络如何实现“动态、安全、可信”的任务协同提供了工程性支撑。但同样,以此为代表,尽管业界已讨论 MCP 与 A2A 的组合可能性,但仍缺乏系统性结构设计。

2.4 未来的智能体:从“互联”走向“互协”

正如互联网依靠 TCP/IP 构建起“万物互联”的基础设施, A2A/MCP 可看作“万智互协 02”的底层协议基座。但仍必须正视一个事实: MCP×A2A 所构建的是一种以任务为中心的功能互联互通体系,而非真正意义上的人与智能体,多人之间通过多智能体的互协生态,人与智能体、人与人之间的隔阂仍然存在。MCP×A2A 的不足并不主要体现在互联与调用能力上,而更集中表现为协作节奏与参与者“身-知-心”3 层结构在不同主体之间的持续失配与累积放大,并不能实现人(团队)与智能体群的共协交互(synergized interaction)^[20-21],使得每个参与者的身体、知识和心理层面状态和发展得到考虑。

在实际的远程协作与多智能体协同场景中,协作问题往往并非源于单一任务失败或技术错误,而是源于协作节奏与参与者状态之间的逐步失配。例如,在持续协作过程中,一方参与者可能因任务负载累积或节奏加快而逐渐出现疲劳迹象,这类变化首先体现在行为与节奏层面(身层);然而,在任务分工与进度安排上,其状态仍被系统或协作方判断为“可持续”(知层);与此同时,其心理状态可能已转向压力增加或心理安全感下降(心层)。当另一方参与者或其智能代理未能及时感知并响应这些变化时,协作往往仍沿既定路径推进,从而放大张力、误解与隐性冲突风险。

基于这一观察,本研究提出的“智能互协网”并非仅关注任务执行层面的优化,而是将协作视为一个跨层、跨主体的动态调节过程。其核心思想在于:通过共协交互机制,在多智能体之间共享关键的状态摘要,并在检测到跨层或跨主体失配时,触发必要的调节介入,从而在协作节奏、任务负担与心理安全之间实现动态再平衡。需要强调的是,这一调节过程并不依赖预设的单一优化目标,而是基于对多层状态的综合判断,决定是否需要介入、介入强度及优先缓解的对象。这一问题意识与协作背景,为后续章节中对智能互协网整体架构以及协作调节协议的形式化描述提供动机基础。

3 基于 HEC×A2A/MCP 的“智能互协网”构想

为回应第 2.4 节中所指出的现有 MCP×A2A 体

系在深度柔性协作方面的不足,引入 HEC 的理念,并在此基础上提出一种新型人机共协网络架构——智能互协网(smart synergetic network, SSN)。该架构并非替代 A2A 与 MCP,而是对其功能的拓展与深化,旨在支持更高层次的状态感知、意图理解与共感反馈。下文将结合 HEC 框架,说明 SSN 的整体结构与设计导向,并进一步讨论“身-知-心”3 层在协作调节中的联动机制及其现实挑战。

3.1 HEC 的意义与定位

HEC 用于刻画人与计算系统在协作过程中形成的理想“共协态”与“共协交互”关系^[20-23]。不同于传统人机交互主要关注任务完成效率,HEC 更强调人在协作过程中的成长性、主体性维护以及价值共创与共感性支持。

如图 1 所示,HEC 由共协人(engaged human)、共协计算机(engaging computer),以及二者之间的共协交互(synergized interaction)构成。为刻画人在协作过程中的多维状态,HEC 理论进一步将共协人抽象为“身-知-心”3 层结构,分别对应行为与生理状态、认知与任务处理能力,以及心理安全、信任与价值取向等因素。该 3 层结构并非对既有理论的简单叠加,而是在综合相关领域研究成果的基础上,对人类经验结构进行更高层次的整合性表达。

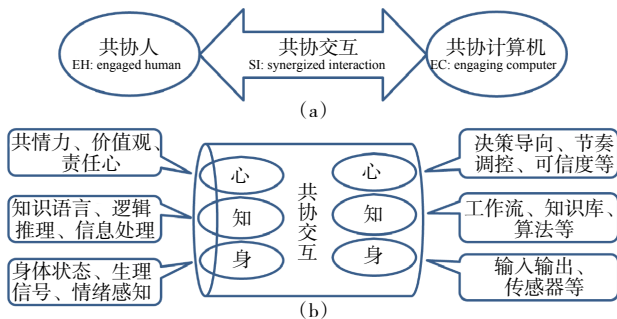


图 1 HEC 模型

与认知架构模型(如 ACT-R)侧重于建立个体内部认知、记忆与意图的模块化结构不同^[24],HEC 的“身-知-心”模型具有更为宏观的理论视角。其中,“知”不仅描述信息加工与任务理解能力,也强调人的主观建构与意义生成过程;“身”不仅涉及感知—行动回路中的物理与生理基础,也将情绪、行为与节奏状态视为认知活动的重要体现;而“心”则进一步指向认

知与情绪背后的价值取向、信任关系与心理安全这一“来源容器”。这种抽象并不与 ACT-R 等认知架构模型相冲突。相对于后者致力于刻画个体层面的认知过程与任务执行机制,HEC 的“身-知-心”模型,用于理解人在协作情境中所呈现的整体经验结构,从而为个体认知研究提供一个更宏观的协作与交互语境基础。

相较于人机共情研究、共情模型,以及情感计算(affective computing)等方向主要聚焦情绪状态的识别、建模、模拟与响应^[25-27],HEC 的“身-知-心”模型在将表层情绪与感受理解为身体化认知反应(“身”)之后,进一步将如何应对情绪的负面性(如焦虑、压力与协作疲劳)作为重要议题,并引入正念、反思与心理调节等理论视角,探讨“心”在情绪平衡、心理安全与关系修复中的调节意义。相比将情绪作为输入特征或优化目标的工程化路径,HEC 更关注情绪状态的生成根源及其在长期协作中的演化逻辑。

此外,与自适应 workflow 系统(adaptive workflow systems)主要根据用户状态或环境变化动态优化任务流程与资源配置不同^[28],HEC 并不止步于“如何调整流程”,而是试图通过讨论行为、认知、情绪、意图与主体性之间的内在联系,理解用户状态变化的动力学根源及其跨层转换机制。这一视角使 HEC 能够超越单一性能指标或局部效率优化,转而关注协作过程中的稳定性、节奏协调与长期可持续性。

从更广泛的研究背景看,情感计算结合机器学习(machine learning, ML)和混合现实(mixed reality, MR)探讨如何识别、解释和响应人类的情绪、情感及感觉的系统性综述^[29]、基于人-AI 交互与以人为中心人工智能(human-centered AI, HCAI)新兴理念的跨学科领域^[30],过去 40 年认知架构领域在感知、注意、动作选择、记忆、学习和推理等核心认知能力方面的机制与方法^[31],以及人机共情在 AI、机器人语境下设计进路的理论架构综述^[32],均为理解人类认知、情绪与价值对齐问题奠定了重要基础。然而,这些研究大多以人-机二元交互为主要对象,尚未将行为状态、认知与情感体验统一纳入多主体、多智能体协作的调节性应用框架之中;更为重要的是,在其着重描述理论与技术框架之外,对于如何提升由认知、情感等人类能力共同组成的人这一关键目的上仍缺乏重视。

在这一意义上,HEC 的“身-知-心”模型可被视为一种关于人类经验结构的宏观理论抽象,其核心价值在于揭示从外向生理到内向心理之间的张力关系、发生根源及可能的调节路径。相应地,共协计算系统在感知、任务管理与适应性调节等层面与人形成对应关系,其目的并非复制人的内部结构,而是动态感知并响应人的状态变化,从而在协作过程中提供恰当支援。正是在这一理论基础之上,本研究后续提出的智能互协网得以将“身-知-心”从个体经验模型进一步

提升为多智能体协作中的跨层调节基础,为协作调节协议的设计提供统一而稳固的抽象支点。

3.2 智能互协网的架构

图 2 展示了 SSN 的整体架构。SSN 的协作调节并非由 A2A 或 MCP 协议本身完成,而是由嵌入用户专属智能体内的“身-知-心”协作调节协议(Arbiter 逻辑角色)统一决策。在该架构中,A2A 用作多智能体之间的通信与状态同步载体,MCP 负责工具、资源与知识的调用支持。

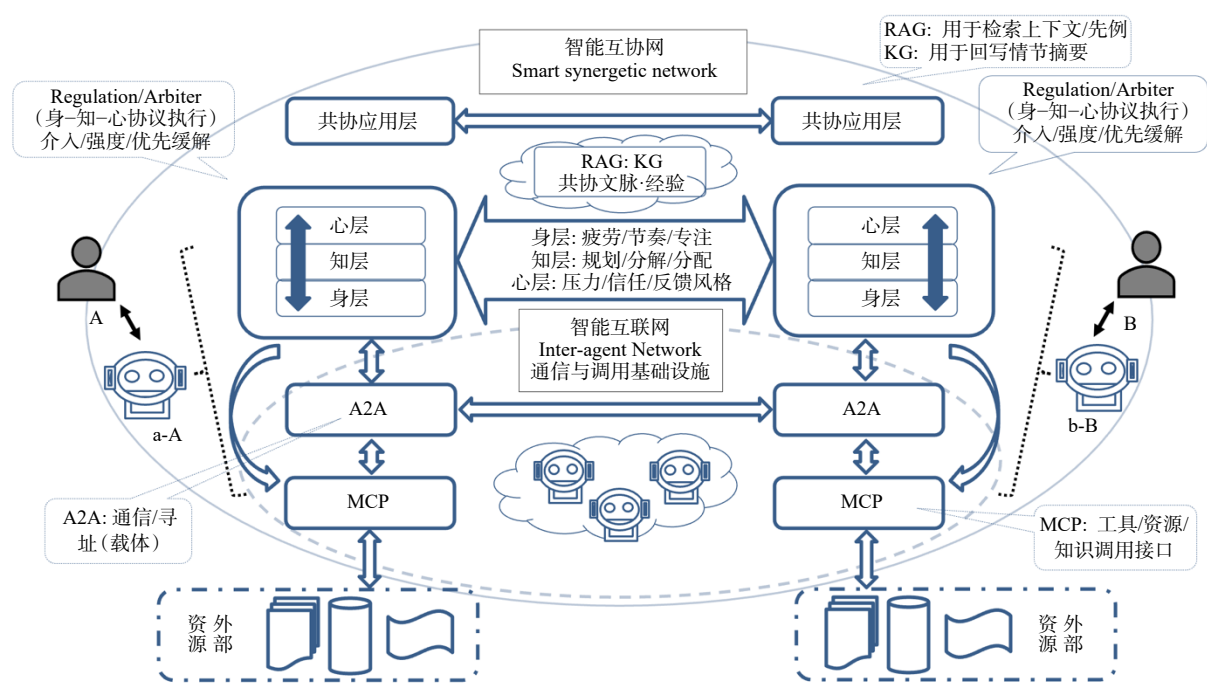


图 2 智协网架构

在 SSN 中,“身-知-心”并非作为抽象的人机模型重复定义,而是作为协作调节协议所使用的 3 类状态维度嵌入代理内部:“身”层提供协作节奏与负荷的行为状态输入,“知”层提供任务结构与进度的认知评估,“心”层提供心理安全与共感相关的调节信号。上述状态在协作过程中被持续汇聚与联动,用于支持是否介入、介入强度以及优先缓解对象的调节判断。除个体用户的专属代理外,SSN 还支持接入云端的支援智能体群组,以在必要时为用户代理提供额外的认知与决策支持。系统亦可结合检索增强生成与知识图谱(knowledge graph, KG)等机制^[33],对协作过程中形成的经验与共识进行管理与应用。上述机制为后续章节中将要介绍的协作调节协议提供了架构基础。

3.3 设计方针:中庸/利他作为协作调节导向的工程定位

智协网的设计目标并非仅在于提升任务执行效率或实现信息层面的互联互通,而是支持一种面向长期协作稳定性与参与者整体状态的调节型协作机制。为此,系统在架构层面引入“身-知-心”3 层结构作为协作调节的输入维度,并在调节决策层明确采用“中庸”与“利他”作为核心设计导向。

在工程实现上,“中庸”与“利他”并不被视为抽象的价值口号,或需要直接数值化的单一优化目标,而是作为协作调节过程中的决策导向(regulation policy orientation),用于约束与引导以下关键判断:是否需要介入协作过程、介入的强度与方式,以及在多方状态

不匹配时优先缓解哪一方的负担与风险。

其中,“中庸”强调调节介入的适度性,即避免过度干预导致协作自主性受损,也避免完全放任造成问题累积,其工程含义体现为对介入阈值、节奏调整幅度与反馈频率的动态控制;“利他”则强调在多参与者协作场景中,调节决策不以单一智能体或单一效率指标的最大化为目标,而是优先关注协作整体的稳定性与心理安全,体现为在任务重分配、节奏调整与反馈策略中对弱势或高风险一方的优先缓解。

因此,诸如信息汇报的频率与粒度、指令提示的强度与时机、协作节奏的调整方式等,并非由系统静态设定,而是结合“身-知-心”3层结构,在调节导向

的约束下进行动态决策。该设计并不试图“控制”协作者行为,而是通过降低摩擦、缓解压力与维持心理安全,为人机共协与多人协作提供一种可持续、柔性的支持机制。上述调节导向在后续章节中通过协作调节协议得到进一步形式化描述。

3.4 基于设计方针的3层联动与协调机制

在智协网中,协议的核心并非单纯的通信或调用规则,而是围绕“身-知-心”3层结构形成的动态协作调节机制。如表1所示,3层各自承担不同类型的状态感知与决策职能,并通过跨层信息传递与反馈,通过跨层联动实现协作稳定性与参与体验的平衡。

表 1 协作流程与“身-知-心”3层结构功能表

协作阶段	身层(行为/状态)	知层(认知/任务)	心层(动机/情感)
状态感知	感知参与者的行为节奏、负载与疲劳状态	识别当前任务结构与依赖关系	识别压力、动机变化与共感需求
协作调节	调整介入时机与协作节奏	进行任务拆分、重分配与角色协调	调整沟通方式,缓解紧张与冲突
稳定反馈	维持可持续的协作节奏	固化有效的协作策略	提供肯定性反馈,增强信任与心理安全

在个体层面,协作调节通常由某一层状态变化触发,但并不固定由“身”层先行。例如,身体负荷或节奏变化可触发认知层对任务节奏与分工的调整;心理压力或安全感下降也可能先被感知,并通过“心”到“知”的反馈路径,引发沟通方式或协作策略的再配置。系统并不假设单一触发源,而是通过多层状态的综合判断,决定是否需要介入以及介入方式。

在多智能体协作中,“知”层通常作为协调枢纽,负责整合来自“身”层与“心”层的状态摘要,用于协商任务分配、节奏同步与协作方式选择;“身”层提供关于负荷与节奏的约束信息;“心”层提供关于动机、压力与心理安全的调节信号,从而避免协作仅围绕任务效率进行单维优化。

在多方或长期协作场景下,3层结构联动进一步体现为一种面向整体稳定性的调节模式:当系统检测到个体或子群体状态失衡时,认知层可触发角色或任务的重新配置,“身”层配合调整行为与节奏,“心”层则通过反馈与共感提示促进团队内部的自然调节与信任恢复。

上述3层结构联动机制并不依赖预设的固定流程,而是在“中庸/利他”调节导向的约束下,根据协作状态动态运行。其具体执行逻辑将在后续的协作调

节协议(见3.6节)中以半形式化方式进一步刻画,并在第3.5节结合知识与经验机制加以扩展。

3.5 智能互协网实现思路

为实现所提出的智能互协网,系统需要在工程层面解决协作状态感知、调节决策与经验积累之间的衔接问题。本研究将其抽象为一个由状态感知—调节决策—执行反馈—经验更新构成的闭环过程,并通过算法给出其半形式化描述。

在状态感知与上报阶段(图3 Step 1),用户专属智能体对“身-知-心”3层结构进行摘要化表示,包括协作节奏、任务负荷、情绪变化与心理安全等指标。这些状态摘要通过A2A作为通信载体在智能体之间同步,并汇聚至仲裁角色(arbiter),为后续调节判断提供输入依据。

在调节决策阶段(图3 Step 2),规则引擎(regulation engine)被定位为跨层调节的决策核心。该模块并非简单执行固定规则,而是在预定义的规则框架内(介入触发规则、介入方式与强度规则、优先级规则),结合当前协作状态,判断是否需要介入、介入强度以及优先缓解对象。“中庸/利他”在此被明确为调节导向,用于约束决策取向,而非作为单一数值化优化目标。

为支持跨时间与跨个体的协作一致性,系统进一步引入基于检索与结构化知识的协作记忆机制。其中,检索增强生成用于从非结构化协作记录、历史案例与对话日志中获取相关上下文信息,知识图谱用于表达任务关系、角色结构与经验规则。二者结合,为算法中的调节判断提供可追溯、可解释的参考依据,用于支持调节决策的背景理解,而非即时响应中的孤立推断。

在执行与反馈阶段(图3 Step 3),调节方案通过A2A作为通信载体在智能体之间实施,包括协作节奏调整、任务重分配与信息呈现方式的适配;在需要时,通过MCP调用外部工具、资源或知识服务作为支撑。协作执行结果被抽象为稳定性、冲突程度、节奏匹配与信任变化等协作结果(OUTCOME)指标,并反馈至系统。

在此基础上,引入强化学习机制,用于在规则结构保持不变的前提下,持续优化调节策略在不同协作情境下的参数配置与组合优先级。Step 3中产生的OUTCOME指标作为奖励信号,用于评估当前调节策略在协作稳定性与心理安全等维度上的综合效果。该奖励信号并不直接驱动即时调节行为,而是通过在线反馈逐步更新调节策略,从而影响后续协作中的调节决策。

与此同时,OUTCOME也被写入RAG/KG,形成可检索、可结构化的协作记忆,用于支持后续协作中的情境参考与经验复用,而不参与即时调节决策。通过上述“执行—反馈—学习—记忆沉淀”的闭环机制,智协网能够在效率之外,逐步学习在不同协作情境下更为合适的调节方式,实现对协作稳定性与心理安全的长期优化,并支撑跨个体、跨时间的人机共协演化。

3.6 智能共协网调节协议算法

根据以上思路,本研究给出智能共协网调节协议算法(伪代码),具体如图3所示。

4 应用场景:智协网的实践可能性

结合“身—知—心”3层结构与A2A/MCP协议机制,提出2个涵盖医疗、教育的潜在应用场景,探索智协网在现实社会中的可行性与应用路径。尽管这些场景处于构想阶段,但初步分析表明,其在人机共协、

```
Participants:
- Agent_A, Agent_B: user-specific agents (biophysics-intellect-xin functions)
- Arbiter: logical mediation role (centralized or distributed)
- Transport:
  - A2A: agent-agent communication carrier
  - MCP: tool / resource / knowledge interface

loop while collaboration_active:

  # Step 1) State sensing and reporting (biophysics-intellect-xin summaries)
  for X in {A, B}:
    sX = summarize_state(X.biophysics, X.intellect, X.xin, context)
    send Arbiter: STATE(X, sX) via A2A

  # Step 2) Arbiter: make Mediation decision (Zhongyong / Altruism as regulation orientation)
  S = aggregate({sA, sB}) # joint evaluation incl. cross-agent mismatch
  R = retrieve_reference(RAG, KG, S) # similar cases, norms, prior patterns (optional)

  plan = null
  if need_intervention(S, R):
    strength = decide_intensity(S, R)
    target = decide_priority(S, R)
    plan = decide_adjustment_plan(S, R) # pacing / redistribution / feedback tuning
    send {Agent_A, Agent_B}: MEDATE(strength, target, plan) via A2A
  else:
    send {Agent_A, Agent_B}: PROCEED(plan) via A2A

  # Step 3) Execution, outcome feedback, and experience update
  for X in {A, B}:
    receive MEDATE or PROCEED from Arbiter via A2A

    execute_via_A2A(X, plan)
    if need_support(plan):
      call_via_MCP(X)

    send Arbiter: OUTCOME(X, outcome_signals) via A2A

  # Arbiter (global learning)
  global_OUTCOME = aggregate (OUTCOME(A, outcome_signal), OUTCOME(B, outcome_signals))
  Arbiter.update_policy(global_OUTCOME)
  Arbiter.update_RAG_KG(global_OUTCOME)

end loop
```

图3 智能共协网调节协议算法

多方协调与共感交互方面展现出一定潜力,为相关产业系统的智能转型提供了值得进一步验证的方向。

4.1 远程医疗团队的协同支持

随着人口老龄化与地区医疗资源分散,医生、护士、康复师等多专业人员之间的远程协作需求日益增加。然而,由于职责差异和工作负荷不同,信息共享的频率和细节层级常常不一致,许多重要却微妙的语境和经验在交流过程中并没有被传递,容易引发误解与延误;同理,这样的问题也发生在医患之间。根据经济学中的“双边信息不对称模型”,促进各方干系人间的信息透明、真诚沟通将为所有人带来最大利益,以及可持续的环境等。

解决方案(“身—知—心”3层结构共协)如下。

“身”层: 监测医护人员的注意力、过劳与睡眠状态, 适时建议轮班或重新分配任务。“知”层: 根据接收者专业背景, 自动调整病历摘要与注意事项的表述粒度(报告频率协议); 鼓励交流各方的信息透明, 保证 A2A 的训练和传播过程能够提供一套完整解释, 减少因语言摩擦、信息不确定性导致的冲突可能性。“心”层: 鼓励真诚合作, 对责任心强而易自责的护士, 通过 A2A 连接生成具有共情力的、团队共同体的安慰性反馈, 缓解心理压力。

该解决方案的特点为, 在保障医疗效率的同时, 调节任务强度, 以利他精神维护以患者为中心的合作模式。

4.2 项目式学习(project-based learning, PBL)的智能支持

跨专业、跨能力的学生共同参与项目型课程, 需要在分工与特点、进度与心理安全感、分心与专注之间取得动态平衡。

解决方案如下。“身”层: 感知每位学生的专注节奏与作业时段, 智能调控小组任务时间点。“知”层: 对经验较少的学生提供个性化补充材料; 对熟练者则细化任务并优化分配(角色调整协议); 智能体能够帮助学生分析其特点和专长, 并且为其分工提供结构化的学习路径与补充视角; 并通过利用 A2A 网络相互沟通, 将团队共协知识库中的各视角进行可视化, 了解各个成员的特长。建立一个清晰的讨论平台, 寻找最合适分工路径, 并帮助寻找团队最合适的组织架构。“心”层: 系统识别沉默学生的非言语信号, 为其提供“代言”, 引导团队成员主动关心与回应(共感路由协议)。

该解决方案的特点为, 尊重个体差异, 通过调整参与深度, 构建利他导向的学习团队。

5 结论与展望

AI 的快速演进, 正持续推动以“智能化”为核心的新一轮工业革命。在这一进程中, 多智能体系统逐渐成为连接人类与 AI 系统、支撑复杂社会场景协同运行的重要基础。A2A 与 MCP 等协议作为当前智能体通信与工具调用的关键机制, 为跨平台协作与外部资源整合提供了统一的技术底座, 构成了智能互联网

的重要基础设施。

在此背景下, 本研究提出了超越既有智能体互联协议的 SSN 架构。通过引入“身-知-心”3 层结构, 以及以利他性与平衡性为导向的调节机制, 从 HEC 的视角, 对多智能体系统中存在的节奏冲突、理解偏差与协作失稳问题给出了系统性的回应, 为未来人机交互协议的设计提供了一种新的结构化范式。

需要强调的是, 智协网的价值并不仅体现在算法或协议层面的创新, 更体现在其所倡导的一种以心理节奏调适、关系修复与共同成长为核心的协作生态观, 以人的提升和向善为目的。这一框架对于教育、医疗、远程协作等多主体参与的复杂场景具有较高的适应性与扩展潜力, 为构建可信赖、可共感的人机共协系统提供了重要的基础。

从工程实现角度看, 以架构与机制提出为主, 尚未展开完整的系统级实现与实验验证。未来工作将围绕在 A2A/MCP 协议之上构建 SSN 展开, 在远程协作与教学等实际场景中逐步探索可获取的数据类型(如任务节奏、交互频率、状态反馈等), 并进一步分析跨层协调、个体差异建模与信任反馈回路更新过程中可能面临的技术瓶颈及其解决策略。

此外, 尽管以 Google 公司为代表提出的 A2A 开源智能互联框架是否会发展为事实标准仍有待观察, 但智能体之间统一通信规范与协同协议的形成已是不可逆的趋势。本研究提出的以“身-知-心”为核心的智协网架构并不依赖于单一协议体系, 其理念与调节机制同样可部署于其他智能体互联框架之上。

展望未来, 如何进一步将利他性原则与中庸理念转化为工程上可操作、可调节的量化体系, 如何在真实系统中持续优化跨层调节机制, 将是智协网走向实际应用的关键研究方向。随着 AI 系统不断融入社会基础结构, 本研究所倡导的“以人机共协为中心”的交互协议设计, 有望为人类与 AI 的共协进化提供长期而稳健的技术支点。

致谢: 本研究受到高知工科大学任向实教授启发和鼓励下, 在内容讨论和修改过程中, 任向实教授提出了宝贵意见。同时, 复旦大学粤港澳大湾区精准医学研究院(广州)檀鹏博士提供了参考资料和修改建议。

参考文献(References)

- [1] 褚君浩. 迎接智能时代[J]. 科技导报, 2025, 43(7): 1–2.
- [2] 吴文峻, 廖星创, 赵金琨. DeepSeek技术创新与通用人工智能发展趋势[J]. 科技导报, 2025, 43(6): 14–20.
- [3] 刘伟, 邹阳洋, 孙维一. 跨越学科、领域、文化和文明的智能[J]. 科技导报, 2024, 42(23): 62–69.
- [4] Metz R. OpenAI Sets Levels to Track Progress Toward Superintelligent AI[EB/OL]. (2024-07-11) [2026-01-03]. <https://www.bloomberg.com/news/articles/2024-07-11/openai-sets-levels-to-track-progress-toward-superintelligent-ai>.
- [5] Google Developers. Announcing the Agent2Agent Protocol (A2A)[EB/OL]. (2025-04-09) [2025-06-15]. <https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/>.
- [6] Anthropic. Introducing the Model Context Protocol [EB/OL]. (2024-11-25) [2025-06-15]. <https://www.anthropic.com/news/model-context-protocol>.
- [7] 网络通信与安全紫金山实验室, 江苏省未来网络创新研究院, 人工智能与数字经济广东省实验室(深圳), 等. 未来网络白皮书: 智能互联网白皮书[EB/OL]. (2022-08-24) [2026-01-03]. <https://aim8.dlssyht.cn/u/551001/ueditor/file/276/55-1001/1662341149630326.pdf>.
- [8] Finin T, Fritzson R, McKay D, et al. KQML as an agent communication language[C]//Proceedings of the Third International Conference on Information and Knowledge Management – CIKM '94. New York: ACM, 1994: 456–463.
- [9] Foundation for Intelligent Physical Agents. FIPA 97 Specification Part 2: Agent Communication Language: FIPA00018 [EB/OL]. [2025-06-15]. <http://www.fipa.org/specs/fipa00018/OC00018A.html>.
- [10] 佐藤, 程子学, 中村, 等. 利用者の心理要素を考慮したエージェントの説得機構を用いた配属支援法[J]. 情報処理学会論文誌, 2002, 43(2): 513–529.
- [11] Ogino T, Kitagami S, Shiratori N, et al. A multi-agent based flexible IoT edge computing architecture and application to ITS[J]. Journal of Communications, 2019: 47–52.
- [12] Guo T C, Chen X Y, Wang Y Q, et al. Large language model based multi-agents: A survey of progress and challenges [EB/OL]. [2025-06-15]. <https://arxiv.org/abs/2402.01680>.
- [13] Langchain Documentation. LangGraph: Multi-agent workflows[EB/OL]. [2025-06-15]. https://langchain-ai.github.io/langgraph/concepts/agent_concepts/.
- [14] Pandya K, Holia M. Automating Customer Service using LangChain: Building custom open-source GPT Chatbot for organizations[EB/OL]. [2025-06-15]. <https://arxiv.org/abs/2310.05421>.
- [15] Barbarroxa R, Gomes L, Vale Z. Benchmarking large language models for multi-agent systems: A comparative analysis of AutoGen, CrewAI, and TaskWeaver[C]//Advances in Practical Applications of Agents, Multi-Agent Systems, and Digital Twins: The PAAMS Collection. Cham: Springer, 2025: 39–48.
- [16] Wu Q Y, Bansal G, Zhang J Y, et al. AutoGen: Enabling next-gen LLM applications via multi-agent conversation [EB/OL]. [2025-06-15]. <https://arxiv.org/abs/2308.08155>.
- [17] Moura J. CrewAI: Fast and flexible multi-agent automation framework[EB/OL]. [2025-06-15]. <https://github.com/crewAIInc/crewAI>.
- [18] Ehtesham A, Singh A, Gupta G K, et al. A survey of agent interoperability protocols: Model Context Protocol (MCP), Agent Communication Protocol (ACP), Agent-to-Agent Protocol (A2A), and Agent Network Protocol (ANP) [EB/OL]. [2025-07-11]. <https://arxiv.org/abs/2505.02279>.
- [19] Jeong C. A study on the MCP×A2A framework for enhancing interoperability of LLM-based autonomous agents [EB/OL]. [2025-07-11]. <https://arxiv.org/abs/2506.01804>.
- [20] 任向实, 付志勇, 麻晓娟, 等. 人机共协计算[M]. 北京: 清华大学出版社, 2024.
- [21] 王晨, 任向实. 从人机交互到人机共协计算: 人机关系的思想演化和未来展望[J]. 科技导报, 2024, 42(8): 6–20.
- [22] Ren X S. Rethinking the relationship between humans and computers[J]. Computer, 2016, 49(8): 104–108.
- [23] Ren X S, Silpasuwanchai C, Cahill J. Human-engaged computing: The future of human-computer interaction[J]. CCF Transactions on Pervasive Computing and Interaction, 2019, 1(1): 47–68.
- [24] Anderson J R, Bothell D, Byrne M D, et al. An integrated theory of the mind[J]. Psychological Review, 2004, 111(4): 1036–1060.
- [25] Gen U, Verma H. Situating empathy in HCI/CSCW: A scoping review[J]. Proceedings of the ACM on Human-Computer Interaction, 2024, 8(2): 1–37.
- [26] Preston S D, de Waal F B M. Empathy: Its ultimate and proximate bases[J]. The Behavioral and Brain Sciences, 2002, 25(1): 1–20.
- [27] Picard R W. Affective computing[M]. Cambridge, MA, USA: The MIT Press, 2000.
- [28] Klein M, Dellarocas C, Bernstein A. Adaptive workflow systems[J]. Special Issue of Computer Supported Cooperative Work, 2000, 9(3/4): 265–267.
- [29] Afzal S, Ali Khan H, Jalil Piran M, et al. A comprehensive survey on affective computing: Challenges, trends, applications, and future directions[J]. IEEE Access, 2024, 12: 96150–96168.
- [30] 许为, 葛列众, 高在峰. 人-AI交互: 实现“以人为中心AI”理念的跨学科新领域[J]. 智能系统学报, 2021, 16(4):

- 605–621.
- [31] Kotseruba I, Tsotsos J K. A review of 40 years of cognitive architecture research: Core cognitive abilities and practical applications[EB/OL]. [2025–07–15]. <https://arxiv.org/abs/1610.08602>.
- [32] 郭景. 情感人工智能时代的用户体验塑造: 人机共情的新路径[J]. 人文社会发展, 2024(1): 13–18.
- [33] Yu H Q, McQuade F. RAG–KG–IL: A multi–agent hybrid framework for reducing hallucinations and enhancing LLM reasoning through RAG and incremental knowledge graph learning integration[EB/OL]. [2025–07–20]. <https://arxiv.org/abs/2503.13514>.

From intelligent agent interconnection to smart synergetic networks

CHENG Zixue^{1,2}, WANG Chen^{2,3}

1. School of Computer Science and Engineering, The University of Aizu, Fukushima 9658580, Japan

2. Shanghai University of Engineering Science, Shanghai 201620, China

3. Kochi University of Technology, Kochi 7828502, Japan

Abstract With the rapid development of Large Language Models (LLMs), communication mechanisms in Multi-Agent Systems (MAS) have evolved from ontology-based protocols such as KQML to LLM-driven architectures. This paper reviews the Agent-to-Agent (A2A) protocol proposed by Google and the Model Context Protocol (MCP) introduced by Anthropic, which together form the MCP×A2A protocol stack. This stack provides a unified interface for agent-level tool invocation and function sharing, forming a foundational layer for the so-called "Internet of Agents". However, the current MCP×A2A system primarily enables functional interoperability and remains insufficient for supporting deep collaboration involving multiple users and agents, especially in terms of dynamic coordination and emotional alignment. To address these limitations, we propose the Smart Synergetic Network, built on A2A/MCP, and as a human-centered architecture grounded in Human-Engaged Computing (HEC). It introduces three integrated layers—Biophysics, Intellect, and Xin—to support full-spectrum collaboration by sensing physical state, structuring task flows, and adjusting emotional dynamics. This layered system dynamically adapts collaboration rhythms and interaction styles to individual user states, enhancing agent understanding and responsiveness. By addressing issues such as mismatched pacing and insufficient empathy in human-AI coordination, the architecture facilitates more flexible, sustainable, and trust-based collaboration. Case studies in education and remote healthcare demonstrate the system's potential to shift AI applications from simple interoperability to deeply synergized interaction.

Keywords human-computer interaction; human-engaged computing; multi-agent system; A2A; MCP; smart synergetic networks ●



(责任编辑 王微)