

AI交通科学家：大模型驱动的自主交通科研

宫晓燕¹, 戴星原^{1,2}, 李芮霖³, 吕宜生^{1,4}

(1. 中国科学院自动化研究所多模态人工智能系统全国重点实验室, 北京 100190;

2. 道路交通安全管控技术国家工程研究中心, 北京 100006;

3. 山东交通学院, 山东 济南 250357;

4. 澳门科技大学创新工程学院, 澳门 999078)

摘要: 城市交通系统正加速演化为社会物理信息系统 (cyber-physical-social system, CPSS), 无人车、无人机及各类智能体持续融入, 导致系统复杂性、动态性与耦合性急剧增强。传统以人工为主的“作坊式”科研模式, 难以实现对系统快速演化行为的深入解析与实时响应。为应对上述挑战, 立足平行智能理论, 提出融合大模型认知能力与多智能体协同优势的自主交通科研框架——AI交通科学家。该框架构建交互层、认知层、实验层、支撑层4层架构, 并以动态路由为调度引擎, 针对机理发现、策略验证与系统优化3类交通科研问题, 通过动态路由自适应编排各层智能体协作链路, 形成匹配不同任务需求的科研流程。框架贯通文献分析、问题识别、假设生成、仿真验证、策略生成、反馈优化的全链条人机协同闭环, 不仅能实现复杂交通问题的自动化识别与高效求解, 还能驱动交通系统规律的自主挖掘, 以及管控策略的创新迭代、多场景验证与持续优化, 为CPSS背景下交通科研的智能化转型提供可落地的技术范式。

关键词: 社会物理信息系统; 平行智能; 自主交通科研; 交通科学家; 自驱实验室

中图分类号: TP39

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202605

AI transportation scientist: LLM-driven autonomous transportation research

Gong Xiaoyan¹, Dai Xingyuan^{1,2}, Li Ruilin³, Lyu Yisheng^{1,4}

1. The State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

2. National Engineering Research Center for Road Traffic Safety Management and Control Technology, Beijing 100006, China

3. Shandong Jiaotong University, Jinan 250357, China

4. Faculty of Innovation Engineering, Macau University of Science and Technology, Macao 999078, China

Abstract: Urban transportation systems are rapidly evolving into CPSS (cyber-physical-social system), driven by the continuous integration of autonomous vehicles, unmanned aerial vehicles, and diverse intelligent agents. This evolution has dramatically increased system complexity, dynamics, and coupling, rendering traditional human-centric research paradigms insufficient for timely understanding and response to fast-evolving system behaviors. To address these challenges, an autonomous framework called “AI Transportation Scientist” was proposed to revolutionize transportation research

收稿日期: 2025-12-16; 修回日期: 2026-03-02

通信作者: 吕宜生, yisheng.lv@ia.ac.cn

基金项目: 国家自然科学基金项目 (No.62271485, No.62303462); 道路交通安全管控技术国家工程研究中心开放课题 (No.2024GCZXKFKT11A); 山东省交通运输厅科技计划项目 (No.2024B70); 山东高速集团有限公司科技计划项目 (No.HS2023B044)

Foundation Items: The National Natural Science Foundation of China (No.62271485, No.62303462), Open Project of the National Engineering Research Center for Road Traffic Safety Management Technology (No.2024GCZXKFKT11A), The Science and Technology Project of Shandong Provincial Department of Transportation (No.2024B70), SDHS Science and Technology Project (No.HS2023B044)

through parallel intelligence. The architecture leveraged a synergy between large language model and multi-agent system across four functional layers (interaction, cognitive, experimental, and support). At its core, a dynamic routing engine adaptively scheduled intelligent agents to tackle mechanism discovery, strategy validation, and system optimization. By implementing a full-chain collaborative closed loop—encompassing problem identification, simulation, and feedback optimization—the framework enabled the autonomous discovery of transportation laws and the continuous evolution of control strategies. This research establishes a scalable technical paradigm for advancing transportation science within CPSS environments, ensuring both efficient problem-solving and innovative strategy iteration.

Key words: CPSS, parallel intelligence, autonomous transportation research, transportation scientist, self-driving laboratory

0 引言

城市交通系统处于一场深刻的结构性转型之中，正从传统的物理系统加速演进为社会物理信息系统（cyber-physical-social system, CPSS）^[1-2]。随着无人车、无人机、路侧智能体、具身机器人以及虚拟交通员工等新型主体的持续融入，交通系统已从单一的物流网络，转变为一个多智能体共生、跨空间耦合的复杂生态系统，呈现出多尺度、强耦合与高度时变的特征^[3]。

在此类系统中，局部扰动可能被迅速放大，诱发全局性系统风险。例如，在低空无人机配送与地面高峰车流叠加的城市场景中，若空地调度算法未能及时响应气象变化或突发需求，就可能引发“空地拥塞共振”现象，导致整体交通系统瘫痪^[4]。防范此类风险，要求科研体系具备从问题发现、机理探索到技术创新的快速闭环能力，这对交通科学的研究速度与智能化水平提出了前所未有的挑战。

长期以来，交通科学依赖建模、仿真与优化来揭示系统机理并制定控制策略，其研究范式正逐步从“模型驱动”向“数据驱动”再向“智能驱动”过渡^[5]。模型驱动方法基于物理方程与理论假设，但在CPSS场景下难以准确刻画异构主体之间的非线性互动与动态反馈机制^[6]；数据驱动方法则借助机器学习与神经网络技术，在预测^[7]与控制^[8]任务中取得了显著进展，但仍依赖人工提出问题与固定研究流程，缺乏自主发现问题与持续学习的能力，难以适应开放环境中知识快速演化以及人机智能共生的新需求。

平行智能理论和人工社会+计算实验+平行执行（artificial society+computational experiment + parallel execution, ACP）方法为构建虚实交互的自动化和智能化研究框架提供了理论基础。大语言模型（large language model, LLM）^[9]与多智能体系统

（multi-agent system, MAS）^[10]则为交通科研范式向智能驱动的转变提供了技术路径。LLM具备语义理解、因果推理与知识融合等认知能力，可作为科研流程的智能中枢^[11]；MAS则通过角色封装与协同机制，实现从问题识别、仿真推演到策略反馈的自组织与自进化流程。

自驱实验室已在化学、材料等领域取得重要突破，能够通过自动化的问题发现、假设提出与实验验证推动规律与机理的发现^[12-13]。然而，交通科学的目标不仅在于揭示规律或机理，更在于确保形成的策略具备可执行性和可验证性。例如，在城市出行管理中，研究者可识别“短途高频通勤是拥堵主因”这一规律，但真正的挑战在于如何将其转化为可实施的差异化收费或激励策略，使其具备技术可行性和社会适应性，从而引导交通参与者出行行为的改变。

为应对上述挑战，基于平行智能理论，本文提出融合大模型与多智能体技术驱动的自主交通科研框架——AI交通科学家。该框架由交互层、认知层、实验层、支撑层及动态路由组成，针对交通研究中的机理发现、策略验证与系统优化3类问题，通过动态路由组织各层智能体形成协作链路，完成不同的任务。该框架覆盖从文献分析、问题识别、假设生成、仿真验证到策略生成与反馈优化的人机协同闭环，不仅能实现问题的自动识别与解决，更能推动规律的自主发现与策略的创新、验证与持续优化。

1 相关工作

本节围绕LLM与智能体系统、基础领域科研自动化以及交通科研自动化3个层面，系统梳理相关理论基础与技术进展，为AI交通科学家提供研究背景与理论依据。

1.1 科研自动化架构: 大模型及智能体系统

LLM与智能体系统构成了科研自动化的“认知核心+行动执行”基础架构。LLM提供语义理解、知识整合与推理生成的认知大脑,智能体则将这种认知能力外化为可执行的科研行为,如调用数据库、调度仿真、撰写报告与优化实验,从而形成具备完整问题求解能力的科研智能体。

在科研认知层面,大语言模型经历了从通用知识理解到科学推理强化的演进。早期的开放模型(如LLaMA^[14])展示了开放领域知识表达与语义推理的潜力,为多领域科学语料的泛化理解奠定了基础。随后,SciBERT^[15]、Galactica^[16]与SciGPT^[17]等科学专用模型通过引入结构化学术语料与因果图谱,强化了学术文本的逻辑解析与因果推理能力,显著提升了科学假设生成与文献溯源的可靠性。进一步地,对话式模型(如LaMDA^[18]与GPT-4^[19])实现了科研流程中的语义交互与多模态知识融合,配合思维链^[20],使模型具备了跨学科问题理解、任务规划与人机协同推理的能力。

在科研执行层面,多智能体系统为科研自动化提供了组织化与协同化的技术路径。Zhang等^[21]提出的多智能体强化学习框架实现了复杂博弈环境中的群体决策支持。随后,Generative Agents^[22]、Auto-GPT^[23]、AgentVerse^[24]与OpenAGI^[25]等框架进一步探索了大模型驱动的智能体在任务分解、工具调用与并行执行中的自治能力。SciAgent系统^[26]将大模型与科学数据库、实验平台相连接,实现了从语义理解到实验设计的初步贯通,但主要停留在认知层闭环,尚未实现实验执行与反馈的自动化迭代。

1.2 自驱实验室: 基础领域科研自动化

自驱实验室是指集成人工智能、机器人技术和自动化实验设备,能够自主完成假设生成、实验设计、实验执行、数据分析的智能化科研平台^[27]。其本质是通过算法闭环替代传统科研中的人工决策环节,实现科研过程自动化和高通量运行。

自驱实验室作为科研自动化在自然科学中的典型实践,已形成从假设生成、实验设计到结果反馈的闭环范式。早期系统(如机器人科学家Adam^[28])实现了基于知识库的自动假设生成与实验验证,然而,这类系统依赖于结构化知识表示,难以适应开放环境中的动态交互与不确定性。此后,自驱动化学实验室^[13]与A-Lab^[29]等项目进一步引入强化学习

与贝叶斯优化方法,实现了实验条件的自主搜索与材料合成的自动化。Wang等^[30]提出的通用自驱动实验室框架进一步融合了大模型与机器人系统,构建了跨学科的科研闭环架构。

尽管自主实验室在化学、材料等基础科学中取得了显著成效,但其实验对象多为物性明确的化学反应或材料体系,实验过程高度结构化。交通系统作为典型的CPSS,其运行过程涉及人类行为、社会规则与动态环境的交织影响,难以直接套用现有自主实验室的封闭式实验范式,亟须发展支持开放环境建模、社会行为仿真与策略闭环验证的新型科研自动化体系。

1.3 交通科研自动化

交通科研自动化旨在构建一个能够自主完成从问题发现到策略优化的智能系统。其发展建立在日趋成熟的理论、平台与认知技术基础之上,正逐步从局部工具自动化向全流程智能驱动迈进。

在交通科研领域,平行智能^[31]为构建虚实互动的自动化研究框架提供了理论支撑。ACP方法^[32-33]阐述了通过构建人工交通系统、开展大规模计算实验并在虚实系统间进行双向引导的方法。在此基础上,平行交通^[34]和并行测试^[35]等系列工作分别面向系统运行管理、智能驾驶测试与科研验证等场景,建立了从模型构建到闭环反馈的技术路径。

在平台层面,数字孪生技术为交通系统的虚拟建模与实验提供了重要工具。在城市规划方面,基于数字孪生的城市交通系统能够实现对路网改造、新建枢纽等规划方案的全生命周期仿真评估^[36-37];在信号优化方面,数字孪生平台通过实时交通流数据驱动信号配比的在线学习与动态调整,显著提升交叉口通行效率^[38-39];在预测及出行诱导方面,借助数字孪生系统对宏观交通状态进行预测^[40],可为交通参与者提供动态、个性化的路径推荐^[41-42];在自动驾驶方面,高保真数字孪生环境为自动驾驶算法的测试验证提供安全、可控且可扩展的虚拟试验场^[43-44]。尽管现有研究在平台构建与局部自动化方面取得显著进展,但其整体科研流程仍依赖人工驱动,尚未实现从问题识别、假设生成到策略优化的全流程自动化闭环。

交通领域大模型正通过其强大的语义理解、知识融合与复杂推理能力,为构建新一代科研智能体提供核心的认知驱动。在交通流预测方面,以UrbanGPT为代表的时空大模型^[45]、基于超图的消息

传递网络^[46]、通用时空预测模型 UniST^[47]以及时空大语言模型^[48]等能够精准捕捉路网中复杂的时空依赖关系,实现对交通状态的多维度精准感知与预测。在信号管控方面, TrafficGPT^[49]和 SignalLLM^[50]探索了自然语言指令与控制策略的映射,基于提示学习的仿真到实景迁移方法^[51]与世界模型框架 TrafficWise^[52]则通过模拟环境动态优化长期控制效益,推动信号控制从经验驱动向认知驱动转变。在交通安全方面, TrafficSafetyGPT^[53]通过微调预训练大模型成为交通安全领域专家, AccidentGPT^[54]构建了面向事故分析与预防的多模态大模型,另有研究利用大语言模型揭示了交通事故漏报问题^[55]。在自动驾驶方面, DriveDreamer^[56]、SurrealDriver^[57]、DriveGPT4^[58]、GPT-Driver^[59]等多模态大模型通过融合视觉、语言与规划,显著提升了系统在开放环境下的场景理解与因果推理能力,世界模型^[60]与图视觉问答方法 DriveLM^[61]进一步增强了自动驾驶系统的环境感知与决策解释性。

综上,平行智能理论为自动化科研奠定了方法论基础,数字孪生平台提供了虚实交互的实验环境,交通领域大模型则赋予了系统理解、推理与决策的专业认知能力。本文旨在将这些要素深度融合,形成一个以大语言模型为通用认知核心、在数字孪生平行系统中实现自我验证与持续演化的科研闭环。

2 AI交通科学家

基于上述分析,本节提出“AI交通科学家”框架。该框架将平行智能理论与大模型驱动的多智能体系统融合,构建了覆盖机理发现、策略验证与系统优化3类科研任务的自动化闭环。

2.1 系统架构

“AI交通科学家”是一个具备自主科研能力的分层架构,由交互层、动态路由、认知层、实验层和支撑层构成。交互层作为人机协同枢纽,动态路由作为自适应调度中枢,认知层承担领域推理与假设生成,实验层负责验证评估与优化,支撑层提供基础设施与模型资源。系统针对机理发现、策略验证与系统优化3类典型的交通科研问题,基于动态路由引擎自适应地从认知层和实验层选择相应智能体,并进行智能体编排,形成相应的科学问题解决链路。系统整体架构及各层协同关系如图1所示,基于“人类在环”的理念,把人类的领域知识、价值判断与伦理监督嵌入自动化流程。

2.1.1 交互层:任务解析与人机协同

交互层是系统与人类科研人员互动的枢纽,由交互智能体构成。其职责有3个方面:一是自然语言解析与任务结构化,即将用户诸如“分析雨季城市突发拥堵成因”的指令,转化为机器可理解、可执行的科研任务图;二是任务调度,将结构化任务图提交给动态路由引擎,启动相应的科研流程;三是结果呈现与交互,将各阶段产出的因果图、评估报告、策略集等成果,以可视化图表和结构化报告的形式反馈给用户,构成透明的双向通信通道。该层为整个自主科研流程提供人机协同及可控可解释的基础。

在“AI交通科学家”框架中,人机协同贯穿全流程,通过预设的关键决策点与开放的随时中断机制,确保人类专家的监督权、指导权与否决权。具体而言,决策点包括以下几点。

在任务启动与校准阶段,系统设立双重保障。一是任务目标确认,交互智能体将解析生成的结构化任务图反馈给用户进行确认与微调;二是伦理与风险预审,系统自动评估任务的社会影响与数据隐私风险,对高风险任务要求人类专家进行伦理审批,锚定研究的合规性与价值导向。

进入认知探索与建模阶段,监督聚焦于科学严谨性。关键因果假设在进入计算实验前需经专家审核其科学性;数据策略与潜在偏见分析报告需经审查,以确保数据基础的可靠性与公平性;复杂模型的可解释性分析需通过专家确认,防止“黑箱”决策,确保模型机理符合领域常识。

在实验验证与优化阶段,协同更为密集且关键。初始策略集的逻辑与潜在风险需经专家评估,专家可删除风险项或注入领域经验;仿真场景的极端性与保真度需经专家确认,以保证压力测试的有效性;评估指标的权重体系可由专家根据实际管理目标动态调整;最终,在多目标优化产生的帕累托最优解集中,必须由人类基于更宏观的社会价值判断进行裁定,确保决策符合综合效益。

此外,系统支持贯穿全程的开放式介入。人类可在任何时刻主动中断自动化流程,进行质询、数据注入、参数调整或任务重定向,系统将完全响应并保持状态可追溯,确保科研活动始终处于人类的可控范围之内。

综上所述,交互层作为系统的人机协同枢纽,其功能为理解人类意图、调度研究任务、呈现与交

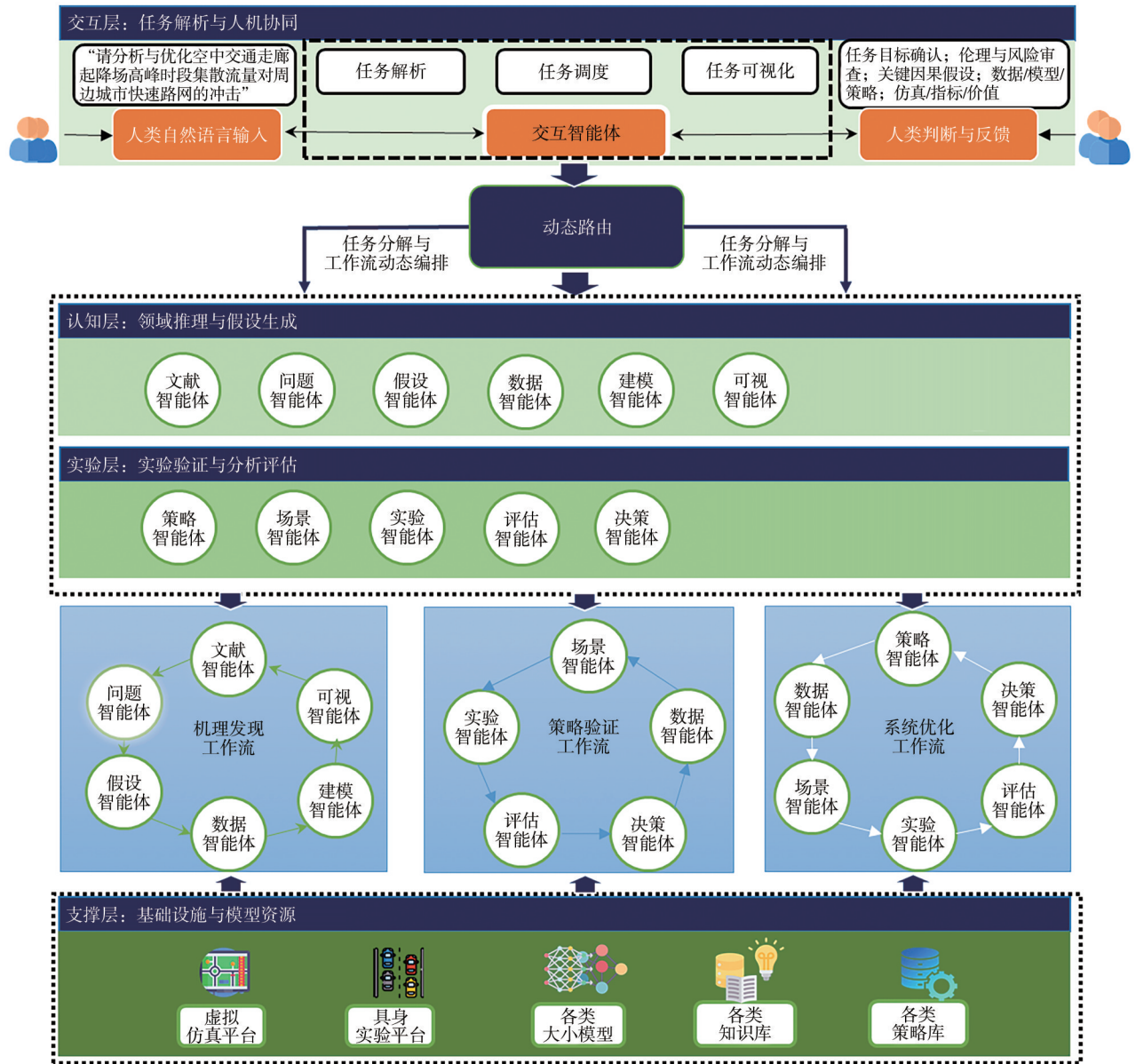


图 1 系统整体架构及各层协同关系

互科研成果，旨在确保人类专家能够以自然的方式启动、引导并理解自动化科研过程。

2.1.2 动态路由：自适应科研工作流引擎

动态路由是架构的调度中枢，承接交互层输出的结构化任务，通过语义解析识别任务类型，并依据智能体能力描述库完成智能体匹配，将交通科研的3类任务（机理发现、策略验证、系统优化）分别映射为认知层与实验层的差异化协作链路，实现科研流程的任务识别、智能体匹配与 workflow 编排，是连接各层级智能体、实现科研流程自适应编排的核心引擎。其中，机理发现面向复杂交通现象的规律识别与理论建模，旨在从多源数据与文献知识中自

动抽取可验证的机理假设；策略验证聚焦于既定理论或控制策略的可行性与有效性检验，目标是实现假设的系统化验证与策略可信度评估；系统优化则针对复杂动态交通系统，通过策略生成、进化与多目标优化，搜索满足多重约束的最优或近似最优解。

(1) 语义解析与任务分解

当科研人员通过交互智能体输入任务，如分析低空无人机对地面拥堵的影响，系统首先利用大语言模型或交通科研大模型进行语义解析，提取任务类型（机理发现、策略验证、系统优化）、研究对象、数据需求及期望输出，解析结果被编码为结构化任务图，输入动态路由引擎。

(2) 智能体匹配与 workflow 编排

动态路由由内置智能体能力描述库, 包含对认知层和实验层中各智能体的输入输出模式、依赖关系与性能标签, 通过任务图与能力图的匹配, 自动生成智能体调用序列。若任务类型为机理发现, 则激活文献智能体→问题智能体→假设智能体→数据智能体→建模智能体→可视智能体的流程; 若任务类型为策略验证, 则激活数据智能体→场景智能体→实验智能体→评估智能体→决策智能体的流程; 若任务类型为系统优化, 则激活数据智能体→策略智能体→场景智能体→实验智能体→评估智能体→决策智能体的流程。

(3) 动态自适应执行与反馈学习

在路由执行过程中, 系统持续监控每个智能体的输出质量与不确定性, 若发现信息不足或结果不收敛, 动态路由引擎会自动调整任务流, 如重新触发数据智能体补充数据或调整建模范式。同时, 路由引擎具备强化学习机制, 通过长期积累不同任务的执行轨迹, 学习最优的智能体组合策略, 实现跨任务迁移与科研流程自进化。

总而言之, 动态路由引擎是“AI 交通科学家”的调度中枢。它通过问题分类、智能体匹配、动态调优的三步机制, 将抽象的科研问题转化为可执行的智能体协作蓝图, 使系统能够灵活、高效地应对机理发现、策略验证与系统优化等不同类型的科研任务, 并具备从经验中学习进化的能力。

2.1.3 认知层: 领域推理与假设生成

认知层承担领域推理与假设生成的职能, 由文献智能体、问题智能体、假设智能体、数据智能体、建模智能体、可视智能体构成, 完成知识整合、问题识别、假设生成、数据融合、模型构建与可视化的过程, 为后续实验验证与策略优化提供科学、可检验的理论起点, 是实现自主科研从知识到假设的核心环节。

(1) 文献智能体

文献智能体负责快速整理科研现状并生成研究语境, 其核心功能是自动化检索多源文献, 利用大语言模型或交通科研大模型的语义解析能力提取研究主题与发展脉络, 并构建跨领域的交通知识图谱。该智能体能够识别交通研究中的知识空白和潜在方向, 为后续智能体提供理论背景和问题线索。

(2) 问题智能体

问题智能体的任务是结合文献与实时及历史交

通数据识别研究问题, 语义化描述现象特征。它基于大语言模型或交通科研大模型对异常现象进行语义化描述, 并将数据异常转化为研究问题。比如, 当问题智能体检测到雨雾天气下无人机低空飞行与地面交通延误同步上升时, 它会将此现象表述为值得进一步探索的科学问题。

(3) 假设智能体

假设智能体针对问题智能体的识别问题提出因果性假设, 并生成形式化可检验命题与因果图。它基于大语言模型或交通科研大模型生成自然语言的研究假设, 经因果推理和数学建模支持, 将假设转化为可计算、可检验的形式。比如, 它可提出“低空无人机高度下降会导致地面交通效率下降”的假设, 并自动生成因果图或方程式。

(4) 数据智能体

数据智能体基于假设智能体生成的假设, 自动在支撑层的数据库中选择合适的数据源、处理方法与特征抽取模型, 为建模智能体提供结构化、可解释的数据输入。它可自动选择并整合实时或者历史的多源交通数据, 包括交通流、气象、视频、通信等, 并进行因果分析与多模态特征抽取和学习, 从数据中提取高维语义特征, 为建模提供数据支撑。

(5) 建模智能体

建模智能体根据数据智能体输出的特征关系构建机理模型或预测模型, 功能包括模型结构搜索、参数估计与模型验证。它支持从传统统计建模到神经符号混合建模的多范式融合, 能够自动匹配任务需求选择合适的模型类型, 如因果图模型、交通动力学模型、深度神经网络模型等, 输出可用于实验验证与策略评估的可计算模型。

(6) 可视智能体

可视智能体实现科研结果的可解释性与形式化表达任务。它基于因果推断与知识图谱可视化技术, 对建模智能体的输出进行语义解释、可视化表达与因果关系说明。可视智能体生成形式化的机理报告或研究简报并发送给交互智能体, 使模型与实验结果能够被科研人员和决策者直观理解, 实现科研透明化与结果可解释化。

概括来说, 认知层构建了系统的领域大脑, 它模拟并加速了科研前期的思维过程。通过智能体协作, 该层可实现从海量、异构的原始信息(文献与数据)到结构化、可检验的科学知识(假设与模型)的自动提炼与升华, 为后续的验证与优化提供

科学起点。

2.1.4 实验层: 实验验证与分析评估

实验层实现假设与策略的生成、验证、评估和优化的任务, 由策略智能体、场景智能体、实验智能体、评估智能体和决策智能体组成, 是实现自主科研从假设到策略的验证环节, 通过虚实融合实验完成策略的有效性、鲁棒性检验与优化。

(1) 策略智能体

策略智能体是实验层的“策略引擎”, 负责将认知层的理论假设转化为具体、可执行的干预方案, 为整个实验流程提供起点和搜索空间。它基于认知层输出的研究假设与明确优化目标, 如“降低核心区域早高峰平均延误15%”, 生成初步的、多样化的策略候选集, 例如特定的信号配时方案、动态收费策略或车辆路径诱导方案。针对同一问题, 它能够生成基于不同作用机理的策略, 如经济杠杆、规则约束、信息诱导, 以探索解决路径的多种可能性。同时, 它为每一个生成的策略附上其内在的因果逻辑链与预期效果的自然语言说明, 使其决策过程透明、可解释。

(2) 场景智能体

场景智能体是虚实融合实验环境的“建筑师”, 可为后续的实验验证生成高保真的虚拟实验场、部署可交互的具身物理实验场及构建虚实融合的实验场。基于数字孪生和生成大模型技术, 该智能体能够动态生成与物理世界实时映射的复杂交通场景, 不仅能复现日常交通流, 更能模拟极端天气、突发交通事故、大型活动散场等极端或特殊工况, 为测试系统的极限承压与应急响应能力提供逼真的虚拟沙盘; 在现实的具身实验场中, 如实验室或封闭测试区等, 该智能体负责调度与部署微型智能车、无人机、机器人警察等具身智能体, 并最终构建与虚拟场景协同联动的物理缩比模型。这些具身智能体既是仿真指令的物理执行终端, 也是真实环境数据的采集节点, 构成了虚实融合的实体锚点。

(3) 实验智能体

实验智能体则是在场景智能体生成的实验环境中开展大规模、高效率实验的“指挥与探索引擎”。它承接来自场景智能体搭建的环境, 并执行虚实实验推演和策略优化等任务。在并行化虚实实验推演方面, 该智能体同时在数字孪生的大规模仿真环境与实体实验场的缩比场景中, 并行执行海量实验任务; 通过在虚拟世界中探索策略的广泛可能

性, 在实体世界中验证关键策略的物理可行性。在策略优化方面, 它驱动虚拟与实体的智能体群体, 运行先进的多智能体协同优化算法, 在复杂的交通场景中自主探索多种管控策略, 如信号配时、路径诱导、无人机巡航等, 并观察策略的联动效应与整体收益。最终, 实验智能体将实验过程数据、性能指标与最终结果传递给评估智能体, 为策略的评估与优化提供数据基础。

(4) 评估智能体

评估智能体负责对实验智能体产生的所有结果进行系统性、多维度的精准评估, 为策略迭代与决策提供权威的量化依据。它可构建一个全面的评估指标体系, 涵盖效率、经济性、安全性与公平性等多个维度, 全面衡量策略的综合效益。然后开展策略对比归因, 不仅判断策略的优劣, 还可识别策略生效的关键因果链条, 最终生成决策报告, 服务于决策智能体的策略优化环节。

(5) 决策智能体

决策智能体是实验层的“决策中枢”与“价值对齐守门员”, 其职责是基于量化评估, 作出可执行的科研决策, 并确保该决策与人类专家的宏观目标、伦理边界和社会价值对齐。它首先对评估智能体提供的多维度量化报告进行深度解读, 不仅要遴选表现最优的策略, 更要分析策略成功的因果机制、潜在的外部性风险以及策略的鲁棒性边界。在此基础上, 决策智能体生成3种初步决策: 采纳(进入部署或知识库)、迭代优化(反馈具体参数给策略/实验智能体)、重启研究(反馈根本性缺陷给认知层)。对于上述初步决策, 尤其是涉及不确定性高、社会影响大或资源消耗多的决策, 决策智能体会把所有初步决策、最优决策、完整的证据链与深度分析, 通过交互智能体提交给人类进行确认与裁决。

综上, 实验层充当了“AI交通科学家”的验证与优化引擎。它将认知层产生的理论构想, 在虚实融合的实验场中通过策略生成、场景构建、实验评估、决策等流程, 转化为经得起检验的实践方案。该层不仅是对假设的测试场, 更是通过演化计算主动探索高性能解决方案的创新场。

2.1.5 支撑层: 基础设施与模型资源

支撑层为整个系统提供稳定、专业的基础运行环境与核心资源, 是“AI交通科学家”实现领域专业化与自主演进的技术底座。该层主要包括虚实

融合实验场、交通领域大模型与小模型以及持续更新的知识库与策略库，是自主科研框架的物理与资源基础，为各层级智能体提供数据、模型、算力与实验环境的支撑。

虚实融合实验场是支撑层的关键基础设施，可将高精度的数字模拟与真实的物理环境紧密结合，形成一个闭环的实验生态系统，包括高保真数字孪生平台和实体具身测试场。高保真数字孪生平台集成高精度地图、物理引擎、交通流仿真模型与丰富的环境模型，如天气、光照等，可构建与物理世界实时同步或预设的极端场景的虚拟镜像，支持大规模、高风险、低成本的并行仿真测试。实体具身测试场则部署于真实世界的封闭或半封闭测试区域，配备真实的微型智能车、无人机、路侧智能单元、机器人警察、交通信号机等可编程实体设备，为虚拟策略提供真实的物理验证环境，收集在真实世界中才存在的噪声、延迟和不确定性数据，确保策略的物理可行性。

另外，支撑层还包括无缝接口与数据桥梁，比如统一的通信协议、数据同步引擎与模型-实体映射系统，确保虚拟空间中的仿真指令能够精确控制实体测试场中的智能体，同时将实体智能体采集的真实数据实时反馈并注入数字孪生平台，驱动虚拟模型迭代优化，形成“虚实互驱”的闭环。

此外，支撑层包含交通领域大模型与小模型和持续更新的数据库、知识库、策略库。交通领域大模型与小模型涵盖交通预测、管控、规划、养护等专业领域，封装了深厚的领域知识与分析方法；持续更新的数据库、知识库与策略库则积累了历史数据、经验与最优解。支撑层通过标准化的应用程序接口，为上层各类智能体提供统一、高效的数据、模型与算力支持。

因此，支撑层是系统的物理底座，可为上层提供统一的数据标准、算力环境及虚实映射接口，使“AI交通科学家”不仅具备认知智能，也具备与物理世界交互的能力。

2.2 3类科研流程

“AI交通科学家”针对机理发现、策略验证与系统优化这3类科研问题，通过动态路由引擎自适应地编排认知层与实验层的智能体，形成3条自动化科研工作流。

2.2.1 机理发现类问题：探索与归纳

该流程旨在探索复杂交通现象背后的内在规律

与因果机制，实现从数据与知识中自动归纳生成可验证的理论假设的目标。工作流由认知层智能体驱动，核心智能体包括交互智能体、文献智能体、问题智能体、假设智能体、数据智能体、建模智能体、可视智能体。

结合“城市突发拥堵形成机理发现”的例子，详细说明其工作流程。科研人员通过自然语言把“分析城市突发拥堵的形成机理”的任务输入交互智能体。交互智能体识别任务类型为机理发现，提取目标变量和期望输出。目标变量包括交通流量、气象条件、信号配时、驾驶行为等，期望输出包括机理模型、因果图谱等。交互智能体最终输出结构化任务描述，包括研究目标、潜在数据类型及智能体调用序列。

文献智能体基于交互智能体输出的任务描述，识别领域关键词，如“突发拥堵”“瓶颈流率”“信号控制”等，开展自动检索和相关文献解析，抽取变量、模型与假设关系，形成交通拥堵知识图谱，最终输出文献驱动的变量网络与研究空白点，如气象与驾驶行为的交互作用尚未形成机理模型。问题智能体基于来自文献智能体的知识图谱与数据智能体提供的实时交通数据，对时空数据进行异常检测和关联分析，识别潜在科研问题，如“雨雾条件下瓶颈流率下降是否触发同步流形成”，输出科学问题陈述与关联变量集。

假设智能体基于科研问题与变量集，利用交通科研大模型的因果推理提出假设，并自动生成因果图结构，如“气象导致瓶颈流率下降→引发同步流→导致系统级拥堵爆发”，最终输出可检验的机理假设及因果模型原型。数据智能体基于假设中涉及的变量关系，调用多源数据，包括传感器、视频、气象、事件记录等，执行特征抽取与因果识别，验证变量间的相关性与时序因果，输出经验证的变量耦合结构与高维特征矩阵。

建模智能体基于数据智能体输出的因果关系与特征矩阵，构建宏微观耦合动力学模型的混合动力学模型，完成参数估计与模型校准，最终输出数据支撑的机理模型与可计算公式。可视智能体基于机理模型结果，执行因果可视化、知识图谱构建与自然语言解释，输出拥堵形成机理的可视化因果链图与关键因子排序。

2.2.2 策略验证类问题：假设与检验

该流程是对已有的理论假设或既定策略进行有

效性、安全性与鲁棒性的系统化验证,包括交互智能体、数据智能体、场景智能体、实验智能体、评估智能体和决策智能体。

本文结合自适应信号控制策略的多场景验证问题,说明策略验证类问题的 workflow。用户通过交互智能体输入任务“验证预测驱动的自适应信号控制在高峰期的有效性”。交互智能体解析任务,识别为策略验证,并将此需求转化为具体、可执行的控制策略候选集,如基于强化学习或模型预测控制的多种信号配时方案。

数据智能体从支撑层的历史数据库中整合相应数据并提供给场景智能体。场景智能体根据验证需求和数据智能体提供的数据,在数字孪生平台中生成多种仿真场景,如常态高峰、事故拥堵、大型活动散场等。实验智能体在这些仿真场景中并行执行所有策略候选,并收集详细的过程数据,如延误、排队长度、通行效率等。

评估智能体对实验数据进行多维度绩效评估与因果归因,量化各策略的优劣。决策智能体基于评估报告对不同策略进行综合权衡,若策略有效,决策智能体可作出采纳决策或提出参数微调建议,并返回交互智能体;若策略未达预期,决策智能体会生成策略修订建议,开启新一轮的验证。最终,流程输出策略的多场景量化评估报告和策略采纳/优化/修订建议。

2.2.3 系统优化类问题: 搜索与迭代

该流程面向复杂动态交通系统,旨在寻找满足多重约束与目标的最优或近似最优解,并通过持续迭代实现策略的进化,包括交互智能体、策略智能体、数据智能体、场景智能体、实验智能体、评估智能体和决策智能体。

以空地协同调度的多智能体优化问题为例,该流程始于用户通过交互智能体输入的自然语言任务,如“优化无人机与地面车辆的协同调度,平衡效率与能耗”,交互智能体随即对任务进行语义解析,将其识别为系统优化类型,并提取关键变量、目标与约束,形成一份结构化的优化任务描述。接下来,策略智能体承接此任务描述,将其转化为一个具体的搜索问题,生成一个多样化的初始策略种群,即多种可能的调度方案,并形式化地定义多目标优化函数,为后续的演化计算奠定基础。

数据智能体从支撑层的历史数据库中整合相应数据并提供给场景智能体。场景智能体根据优化问

题与策略种群的特征和相应数据,在数字孪生平台中构建与之匹配的高保真仿真实验环境,为策略评估提供可靠的测试场。环境就绪后,实验智能体作为核心的“演化引擎”开始工作,驱动策略种群在虚实交互的仿真环境中进行大规模并行实验,并利用进化算法或强化学习等技术,使策略不断地经历评估、选择、交叉与变异的迭代循环,从而自主地演化出性能逐代提升的新策略。评估智能体则在此过程中密切监控,对每一代策略的表现进行多维度量化评估,并生成关键的帕累托前沿,以清晰揭示不同优化目标(如效率与能耗)之间的内在权衡关系。

最后,决策智能体基于评估结果与帕累托前沿行使最终决策权。它综合系统约束与决策偏好,判断当前策略种群是否已收敛到满意解。若已满足要求,则从中选定最终的策略方案输出并返回给交互智能体;若未满足,则自动分析性能瓶颈,调整优化方向或目标权重,并将修订指令反馈给策略智能体,从而开启新一轮“策略生成-实验-评估”迭代。流程最终输出经演化优化的最优策略方案、帕累托前沿分析报告以及策略控制参数集。

2.2.4 概念验证: 贯穿3类流程

为说明“AI交通科学家”框架处理复杂 CPSS 问题的全流程能力与人机协同,本文设计了城市空中交通走廊(urban air mobility corridor, UAM)与地面主干路网动态耦合优化的概念验证案例。此案例将依次触发机理发现、策略验证与系统优化3类科研流程,并贯穿人机协作。

案例如下:在未来特大型城市的空中交通走廊中,电动垂直起降飞行器(electric vertical take-off and landing, eVTOL)与地面车流产生交互。人类管理者向系统提出任务:“请分析与优化空中交通走廊起降场高峰时段集散流量对周边城市快速路网的冲击,并预防系统性瘫痪风险”。

人类管理者通过交互智能体以自然语言下达上述指令,动态路由将其识别为机理发现类问题,并激活机理发现科研链路。系统完成从跨域文献综述、空-地多源数据关联分析到科学假设生成的过程,并提出因果假设“恶劣天气导致eVTOL进近航段效率陡降,引发的航班积压与改航备降需求,在短时间内转化为前往备用起降场的地面保障车辆与旅客接送车流,是触发关键枢纽周边路网级联性拥堵的共振诱因”。将上述假设及支撑的时空关联

数据提交给人类专家审核。人类专家确认其合理性，补充关键的社会行为参数，修正模型中关于旅客在延误场景下对网约车偏好的微观行为偏差。“旅客在面临航班取消或长时间延误时的焦虑情绪，会显著提高其放弃公共交通转而选择即时性更高的网约车出行的概率”。由此，系统构建了融合专家反馈的包含空域流量、地面路网、气象干扰及旅客行为的模型。

基于上述模型，人类管理者下达了新的具体指令“验证基于气象预警的空-地动态流量协同疏导策略的有效性，具体来说，当气象雷达预测未来30 min内UAM将出现效率衰减时，系统自动生成并仿真一组联合指令，包括适当调减起降场中eVTOL的起降架次、向空中航班发布预先改航建议、同步对地面关联快速路匝道实施动态流量诱导与信号协调”。动态路由识别其为策略验证问题，并激活相应科研链路。在由实验层智能体构建的城市级数字孪生环境中，对该策略进行大规模、多场景的仿真测试与量化评估。决策智能体生成的评估报告显示，该策略在模拟的雷暴天气场景下，能平均降低关联区域40%的极端交通延误，避免路网瘫痪。报告同时指出，该策略会导致UAM运营商在高峰时段损失约15%的运营收益。系统将包含公众出行效率提升与商业运营收益损失的量化评估报告提交给人类管理者。管理者统筹之后，给出指令“采纳该方案作为高级别预警下的应急响应机制，但需启动精细化协同优化”。

由此，任务目标成为在多重约束下寻找空-地系统全局更优的常态化运行方案，系统经动态路由识别后进入系统优化阶段。策略智能体生成一个包含大量可变参数的策略种群，如精细化eVTOL时隙分配、地面路网自适应信号配时、对旅客的多模式出行建议等。实验智能体作为演化引擎，驱动该策略种群在虚实同构的具身实验场中进行自主迭代与进化。评估智能体持续计算“整体通行效率”“UAM运营准点率”“社会总出行时间成本”等多个目标的帕累托前沿。当系统演化出一组或者几组最优方案时，决策智能体将最终的判断权交还给人类管理者，并将完整的帕累托分析报告、各方案对应的具体技术参数、预期效果和潜在影响均呈交给人类管理者，由人类管理者依据政策导向选定最终实施方案。

综上，该概念验证案例表明，“AI交通科学

家”框架可将一个涉及未来技术、跨管理域、充满不确定性的复杂系统挑战，系统分解并依次通过机理发现、策略验证、系统优化的科研链条进行处理，形成从理论认知到优化策略的闭环。它展示了“人类在环”协同机制：人类在跨域问题定义、社会行为参数注入、多维利益权衡与重大价值裁决等节点行使监督与指导权，系统则承担了海量异构数据分析、高通量跨域仿真推演与高维策略空间搜索等繁重的计算任务，为人机协同决策提供了支持。

3 应用场景

“AI交通科学家”在行为认知与机理重建、跨域协同与系统最优、开放科研与共治机制等方面具备应用潜力。

3.1 行为认知与机理重建

该场景对应于科研流程中的探索与归纳阶段，聚焦交通CPSS中人机混合环境的行为机制与群体动力学规律识别。传统交通研究多以宏观流量或控制变量为中心，难以揭示社会行为与系统演化的内在机理。“AI交通科学家”通过文献智能体与数据智能体的双源推理，整合行为科学、博弈论与交通动力学知识，能够自动发现并建模复杂的交互规律。

例如，在人类驾驶与无人车混行的城市场景中，系统首先通过文献智能体整理已有研究脉络、提炼行为冲突模式；其次通过问题智能体从真实交通数据中识别异常现象，如人类驾驶的不确定性导致自动车队稳定性下降；再次通过假设智能体提出因果假设“人类随机博弈行为破坏车队同步稳定”；最后通过建模与可视智能体进一步形成可计算的机理模型与可视化因果图谱。这一过程实现了从海量数据到因果机制的自动抽象与知识重构，为理解人机共融交通系统提供理论依据，也为策略验证和优化提供逻辑起点。

3.2 跨域协同与系统最优

该场景对应于科研流程的搜索与迭代阶段，体现“AI交通科学家”在物理与信息空间的协同优化能力。在空地协同、低空经济与智能网联等多域场景中，交通系统的优化问题涉及多目标、多约束与动态交互，传统模型难以实现全局最优。

“AI交通科学家”通过多智能体协同机制，将复杂问题分解为可并行求解的子任务。策略智能体构建初始策略群体，进化智能体在虚实融合仿真环境中执行遗传与强化学习算法，实现策略的自适应

演化; 评估智能体动态计算能耗、延误、安全性等指标; 协调智能体基于帕累托分析实现多目标均衡; 决策智能体最终输出全局最优策略。以空地协同调度为例, 系统可同时优化无人机起降窗口与地面信号周期, 在不同天气与流量条件下保持整体运行效率最优。这一过程展示了“AI 交通科学家”在复杂系统层面实现自演化优化的能力, 为未来城市级交通管控提供可持续的技术路径。

3.3 开放科研与共治机制

不同于前两类以机理探索与策略优化为中心的场景, 本场景聚焦科研范式层的社会化演进。随着去中心化自治组织 (distributed autonomous organization, DAO)^[62]、分布式科学^[63]与去中心化治理理念^[64]的发展, 交通科研正从封闭实验室走向开放协作与社会共治。“AI 交通科学家”可作为开放的科研基础设施与 DAO 机制深度耦合, 支持科研任务的分布式发布、验证与激励。

在该模式下, 文献与问题智能体自动识别公共交通治理中的关键问题, 如学校区域安全、应急交通响应等, 并形成可执行的研究任务。DAO 社区通过代币或信誉机制调动研究者与公众参与, 交互智能体则作为人机接口, 使参与者能提交方案、上传数据或审阅结果。系统内部的仿真实验与评估智能体负责验证与因果分析, 保证科研的科学性与可追溯性。这种机制使交通科研与治理过程实现从专家决策到群体共治的转型, 推动科研民主化、透明化与持续创新。

4 挑战与发展方向

尽管“AI 交通科学家”已构建起从认知生成、策略创新到系统优化的科研闭环, 并展示出交通科研自动化的巨大潜力, 但要真正实现一个可信、可解释、可治理的自主科研体系, 仍需跨越若干关键瓶颈。这些挑战不仅涉及算法与计算, 更触及智能科研的知识结构、伦理规范与社会协同机制。

4.1 复杂系统的因果发现与可解释性瓶颈

交通 CPSS 具有显著的多层耦合与非线性特征, 社会行为、物理约束与信息交互之间形成高度复杂的反馈回路。当前大模型虽然在相关性识别与预测方面表现优异, 但在因果机制辨识上仍存在“幻觉”与“黑箱”问题, 导致科研推理难以被验证与解释。未来亟须突破的方向是从相关性分析走向可验证的因果推断。这要求在“AI 交通科学家”

的认知层中深度融合因果发现算法、交通机理建模与可解释人工智能, 构建具备交通语义约束的推理框架, 使智能体不仅能回答“是什么相关”, 更能阐释“为什么如此”, 为科研智能体提供从观测学习到机理建模的真正科学能力。

4.2 领域知识嵌入与交通领域大模型的构建

通用大语言模型虽具备强大的语言理解与生成能力, 但在交通科研领域常出现常识化理解与非物理推理等问题, 缺乏工程深度。AI 交通科学家的核心挑战之一是如何让模型具备交通工程专家的知识体系与推理范式。科研系统需要实现知识与模型的深度融合, 一方面, 通过知识图谱与符号推理结构化交通流理论、控制原理与政策规范; 另一方面, 将这些符号知识与神经网络语义空间融合, 构建可计算的交通知识图谱+可学习的大模型双支撑架构。未来应发展具备专业约束的交通领域大模型, 实现自然语言理解、物理计算与控制推理的无缝衔接, 从根本上提升科研自动化产出的专业性与可靠性。

4.3 科研智能体的自主性、安全性与伦理控制

在“AI 交通科学家”中, 各科研智能体具备一定的自主生成与执行能力。然而, 这种自主性若缺乏约束, 可能带来安全与伦理风险。例如, 生成未经验证的高风险实验假设、在开放环境中泄露敏感数据或因模型偏差导致不公正决策。未来必须在自主智能与伦理监督之间建立动态平衡机制。技术层面, 应在动态路由与决策智能体中嵌入可信 AI、合规执行与行为审计机制, 确保科研任务在安全边界内运行; 制度层面, 应构建“人类科学家在环”的监督体系, 明确科研责任边界, 防止智能体在自动化过程中偏离科学伦理。通过多层次的安全、法律与社会约束, 为“AI 交通科学家”的长期可信运行奠定基础。

4.4 动态智能体社会的协同与治理难题

随着科研任务复杂度提升, “AI 交通科学家”的智能体将逐步从固定模块演化为可自主重组、动态协商的科研智能体社会。在应对突发拥堵、极端天气或多域协同等场景时, 系统可能临时组建由应急调度、风险评估、优化决策等智能体构成的自治联盟。这种动态自治结构虽提升了科研效率与自适应性, 但也带来了新的治理难题: 如何设计高效的通信与协商机制以保证智能体联盟稳定, 如何在多目标情境下通过机制设计防止个体目标偏离全局任

务, 如何确保群体智能的演化方向与人类价值和伦理保持一致。未来的研究需融合博弈论、组织理论与社会计算等多学科方法, 建立可治理的智能体社会结构, 实现科研智能体间的协调共识与价值对齐。

4.5 平行科研生态的构建与开放协同机制

“AI交通科学家”的最终演化, 不仅是算法体系的升级, 更是科研生态的重构。当前交通科研仍以机构和实验室为中心, 数据孤岛与模型封闭限制了自动化系统的自我进化。未来应依托“平行科学”理念^[65-66], 构建虚实融合、开放协同的科研生态系统。在人工系统中, 科研智能体通过计算实验、虚实比对与多智能体协同, 不断生成、验证与修正假设; 在物理系统和人工系统之间的平行执行中, 借助DAO机制实现科研任务的开放分发、众包参与和社会激励。通过“人工系统、计算实验、平行执行”的人机物三元机制, 科研从封闭系统转向群体智能驱动, 实现知识生产的规模化、透明化与社会化。平行科研生态的形成, 将推动“AI交通科学家”从单一系统走向开放科学共同体, 成为智能社会基础设施的重要组成部分。

5 结束语

本文针对CPSS背景下交通科研的范式变革需求, 提出了“AI交通科学家”这一自主科研框架。该框架融合了平行智能理论、大语言模型的认知能力以及多智能体系统的协同执行力, 通过交互层、动态路由、认知层、实验层、支撑层的分层架构与动态路由机制, 旨在构建覆盖机理发现、策略验证与系统优化的自动化研究链路, 提升科研自主化及效率。

“AI交通科学家”实现了交通科研从人工主导向人机协同的范式演进, 通过交互层提供人机协同及可控可解释的基础, 通过认知层完成从知识到假设的理论提炼, 通过实验层实现从假设到策略的验证优化, 通过支撑层提供虚实融合的实验与资源基础, 全程以“人类在环”机制保障科研的可信性与可解释性。概念验证案例初步说明了框架处理复杂CPSS问题的可行性, 其应用场景涵盖行为认知与机理重建、跨域协同与系统优化、开放科研与共治机制, 为交通科研自动化提供了可落地的技术路径。

参考文献:

- [1] Wang F Y. The emergence of intelligent enterprises: from CPS to CPSS[J]. *IEEE Intelligent Systems*, 2010, 25(4): 85-88.
- [2] 杨静, 王晓, 王雨桐, 等. 平行智能与CPSS: 三十年发展的回顾与展望[J]. *自动化学报*, 2023, 49(3): 614-634.
Yang J, Wang X, Wang Y T, et al. Parallel intelligence and CPSS in 30 years: an ACP approach[J]. *Acta Automatica Sinica*, 2023, 49(3): 614-634.
- [3] Feng Z H, Xue R Q, Yuan L, et al. Multi-agent embodied AI: advances and future directions[PP]. V2. (2025-06-21)[2025-12-16]. arXiv: arXiv.2505.05108.
- [4] He Y X, Wang D W, Huang F H, et al. Aerial-ground integrated vehicular networks: a UAV-vehicle collaboration perspective[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, 25(6): 5154-5169.
- [5] 王飞跃, 缪青海. 人工智能驱动的科学新范式: 从AI4S到智能科学[J]. *中国科学院院刊*, 2023, 38(4): 536-540.
Wang F Y, Miao Q H. Novel paradigm for AI-driven scientific research: from AI4S to intelligent science[J]. *Bulletin of Chinese Academy of Sciences*, 2023, 38(4): 536-540.
- [6] Wang F Y. Parallel control and management for intelligent transportation systems: concepts, architectures, and applications[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2010, 11(3): 630-638.
- [7] Lv Y S, Duan Y J, Kang W W, et al. Traffic flow prediction with big data: a deep learning approach[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(2): 865-873.
- [8] Li L, Lv Y S, Wang F Y. Traffic signal timing via deep reinforcement learning[J]. *IEEE/CAA Journal of Automatica Sinica*, 2016, 3(3): 247-254.
- [9] Radford A, Narasimhan K, Salimans T, et al. Improving language understanding by generative pre-training[EB]. (2018-06-11) [2025-12-16].
- [10] Wooldridge M J. An introduction to multiagent systems[M]. 2nd ed. Chichester: John Wiley & Sons, 2009.
- [11] Brown T B, Mann B, Ryder N, et al. Language models are few-shot learners[C]//*Proceedings of the 34th International Conference on Neural Information Processing Systems*. New York: ACM, 2020: 1877-1901.
- [12] King R D, Whelan K E, Jones F M, et al. Functional genomic hypothesis generation and experimentation by a robot scientist[J]. *Nature*, 2004, 427(6971): 247-252.
- [13] Abolhasani M, Kumacheva E. The rise of self-driving labs in chemical and materials sciences[J]. *Nature Synthesis*, 2023, 2(6): 483-492.
- [14] Touvron H, Lavril T, Izacard G, et al. LLaMA: open and efficient foundation language models[PP]. V1. (2023-02-27) [2025-12-16]. arXiv: arXiv.2302.13971.
- [15] Beltagy I, Lo K, Cohan A. SciBERT: a pretrained language model for scientific text[C]//*Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Stroudsburg: ACL, 2019: 3613-3618.
- [16] Taylor R, Kardas M, Cucurull G, et al. Galactica: a large language model for science[PP]. V1. (2022-11-16)[2025-12-16]. arXiv: arXiv.2211.09085.
- [17] She F, Wang N, Wu H, et al. SciGPT: a large language model for scientific literature understanding and knowledge discovery[PP]. V1. (2025-09-09) [2025-12-16]. arXiv: arXiv.2509.08032.
- [18] Thoppilan R, De Freitas D, Hall J, et al. LaMDA: language models for dialog applications[PP]. V3. (2022-02-10)[2025-12-16]. arXiv: arXiv.2201.08239.
- [19] OpenAI. GPT-4 technical report[R]. 2023.
- [20] Bosma M, Chi E, Ichter B, et al. Chain-of-thought prompting elicits

- reasoning in large language models[C]//Proceedings of the Advances in Neural Information Processing Systems. Massachusetts: MIT Press, 2022: 24824-24837.
- [21] Zhang K Q, Yang Z R, Başar T. Multi-agent reinforcement learning: a selective overview of theories and algorithms[PP]. V2. (2019-11-24) [2025-12-16]. arXiv: arXiv.1911.10635.
- [22] Park J S, O'Brien J, Cai C J, et al. Generative agents: interactive simulators of human behavior[C]//Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology. New York: ACM, 2023: 1-22.
- [23] Richards B. Auto-GPT: An autonomous GPT-4 experiment[EB]. (2023-03-30) [2025-12-16].
- [24] Chen W Z, Su Y S, Zuo J W, et al. AgentVerse: facilitating multi-agent collaboration and exploring emergent behaviors[PP]. V3. (2023-10-23) [2025-12-16]. arXiv: arXiv.2308.10848.
- [25] Ge Y Q, Hua W Y, Ji J C, et al. OpenAGI: when LLM meets domain experts[C]//Proceedings of the Advances in Neural Information Processing Systems. Massachusetts: MIT Press, 2023: 5539-5568.
- [26] Li X, Wu R, Liu X, et al. SciAgent: A Unified Multi-Agent System for Generalistic Scientific Reasoning[PP]. V2. (2025-11-07) [2025-12-16], arXiv: arXiv.2511.08151.
- [27] Leong S X, Griesbach C E, Zhang R, et al. Steering towards safe self-driving laboratories[J]. Nature Reviews Chemistry, 2025, 9(10): 707-722.
- [28] King R D, Rowland J, Oliver S G, et al. The automation of science[J]. Science, 2009, 324(5923): 85-89.
- [29] Symonds J, Thakkar A, Borsos Z, et al. A-Lab: an autonomous laboratory for the accelerated discovery of new materials[J]. Nature, 2023, 624: 86-91.
- [30] Wang H C, Fu T F, Du Y Q, et al. Scientific discovery in the age of artificial intelligence[J]. Nature, 2023, 620(7972): 47-60.
- [31] 王飞跃. 平行系统方法与复杂系统的管理和控制[J]. 控制与决策, 2004, 19(5): 485-489, 514.
Wang F Y. Parallel system methods for management and control of complex systems[J]. Control and Decision, 2004, 19(5): 485-489, 514.
- [32] 王飞跃. 关于复杂系统的建模、分析、控制和管理[J]. 复杂系统与复杂性科学, 2006, 3(2): 26-34.
Wang F Y. On the modeling, analysis, control and management of complex systems[J]. Complex Systems and Complexity Science, 2006, 3(2): 26-34.
- [33] 王飞跃. 计算实验方法与复杂系统行为分析和决策评估[J]. 系统仿真学报, 2004, 16(5): 893-897.
Wang F Y. Computational experiments for behavior analysis and decision evaluation of complex systems[J]. Journal of System Simulation, 2004, 16(5): 893-897.
- [34] 吕宜生, 陈圆圆, 金峻臣, 等. 平行交通: 虚实互动的智能交通管理与控制[J]. 智能科学与技术学报, 2019, 1(1): 21-33.
Lyu Y S, Chen Y Y, Jin J C, et al. Parallel transportation: virtual-real interaction for intelligent traffic management and control[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(1): 21-33.
- [35] 王晓, 张翔宇, 周锐, 等. 基于平行测试的认知自动驾驶智能架构研究[J]. 自动化学报, 2024, 50(2): 356-371.
Wang X, Zhang X Y, Zhou R, et al. An intelligent architecture for cognitive autonomous driving based on parallel testing[J]. Acta Automatica Sinica, 2024, 50(2): 356-371.
- [36] Batty M. Digital twins[J]. Environment and Planning B: Urban Analytics and City Science, 2018, 45(5): 817-820.
- [37] Batty M. Digital twins in city planning[J]. Nature Computational Science, 2024, 4(3): 192-199.
- [38] Liu H Q, Su W K, Li T, et al. Digital twin enhanced multi-agent reinforcement learning for large-scale mobile network coverage optimization[J]. ACM Transactions on Knowledge Discovery from Data, 2025, 19(1): 1-23.
- [39] Dasgupta S, Rahman M, Jon S. Harnessing digital twin technology for adaptive traffic signal control: improving signalized intersection performance and user satisfaction[J]. IEEE Internet of Things Journal, 2024, 11(22): 36596-36618.
- [40] Liao X W, Leng S P, Sun Y, et al. A digital-twin-based traffic guidance scheme for autonomous driving[J]. IEEE Internet of Things Journal, 2024, 11(22): 36829-36840.
- [41] Nie L S, Wang X J, Zhao Q L, et al. Digital twin for transportation big data: a reinforcement learning-based network traffic prediction approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(1): 896-906.
- [42] Thonhofer E, Sigl S, Fischer M, et al. Infrastructure-based digital twins for cooperative, connected, automated driving and smart road services[J]. IEEE Open Journal of Intelligent Transportation Systems, 2023, 4: 311-324.
- [43] Dosovitskiy A, Ros G, Codevilla F, et al. CARLA: an open urban driving simulator[C]//Conference on Robot Learning. New York: PMLR, 2017: 1-16.
- [44] Shah S, Dey D, Lovett C, et al. AirSim: high-fidelity visual and physical simulation for autonomous vehicles[M]//Field and Service Robotics. Cham: Springer International Publishing, 2017: 621-635.
- [45] Li Z H, Xia L H, Tang J B, et al. UrbanGPT: spatio-temporal large language models[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2024: 5351-5362.
- [46] Wang J C, Zhang Y, Hu Y L, et al. Large-scale traffic prediction with hierarchical hypergraph message passing networks[J]. IEEE Transactions on Computational Social Systems, 2024, 11(6): 7103-7113.
- [47] Yuan Y, Ding J T, Feng J, et al. UniST: a prompt-empowered universal model for urban spatio-temporal prediction[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2024: 4095-4106.
- [48] Liu C X, Yang S, Xu Q X, et al. Spatial-temporal large language model for traffic prediction[C]//Proceedings of the 2024 25th IEEE International Conference on Mobile Data Management (MDM). Piscataway: IEEE Press, 2024: 31-40.
- [49] Zhang S Y, Fu D C, Liang W Z, et al. TrafficGPT: Viewing, processing and interacting with traffic foundation models[J]. Transport Policy, 2024, 150: 95-105.
- [50] Ke J L, Hu Q Y, Yuan S H, et al. SignalLLM: a general-purpose LLM agent framework for automated signal processing[PP]. V2. (2025-10-30) [2025-12-16]. arXiv: arXiv.2509.17197.
- [51] Da L C, Gao M Q, Mei H, et al. Prompt to transfer: sim-to-real transfer for traffic signal control with prompt learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(1): 82-90.
- [52] Hu J J, Dai X Y, Li X J, et al. TrafficWise: leveraging world models for generalized and interpretable traffic control[J]. IEEE Intelligent Transportation Systems Magazine, 2025, 17(5): 50-60.
- [53] Zheng O, Abdel-Aty M, Wang D D, et al. TrafficSafetyGPT: tuning a pre-trained large language model to a domain-specific expert in transportation safety[PP]. V1. (2023-07-28)[2025-12-16]. arXiv: arXiv.2307.15311.
- [54] Wang L N, Ren Y L, Jiang H, et al. AccidentGPT: a V2X environmen-

tal perception multi-modal large model for accident analysis and prevention[C]//Proceedings of the 2024 IEEE Intelligent Vehicles Symposium (IV). Piscataway: IEEE Press, 2024: 472-477.

[55] Artega C, Park J. A large language model framework to uncover under-reporting in traffic crashes[J]. Journal of Safety Research, 2025, 92: 1-13.

[56] Wang X F, Zhu Z, Huang G, et al. DriveDreamer: towards real-world-drive world models forAutonomous driving[C]//Computer Vision - ECCV 2024. Cham: Springer, 2025: 55-72.

[57] Jin Y, Yang R X, Yi Z J, et al. SurrealDriver: designing LLM-powered generative driver agent framework based on human drivers' driving-thinking data[C]//Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE Press, 2024: 966-971.

[58] Xu Z H, Zhang Y J, Xie E Z, et al. DriveGPT4: interpretable end-to-end autonomous driving via large language model[J]. IEEE Robotics and Automation Letters, 2024, 9(10): 8186-8193.

[59] Mao J G, Qian Y X, Ye J J, et al. GPT-Driver: learning to drive with GPT[PP]. V3. (2023-12-05)[2025-12-16]. arXiv: arXiv.2310.01415.

[60] Wang Y Q, He J W, Fan L, et al. Driving into the future: multiview visual forecasting and planning with world model for autonomous driving[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2024: 14749-14759.

[61] Sima C, Renz K, Chitta K, et al. DriveLM: driving withGraph visual question answering[C]//Computer Vision - ECCV 2024. Cham: Springer, 2025: 256-274.

[62] Li J J, Qin R, Wang F Y. The future of management: DAO to smart organizations and intelligent operations[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2023, 53(6): 3389-3399.

[63] Ding W W, Hou J C, Li J J, et al. DeSci based on Web3 and DAO: a comprehensive overview and reference model[J]. IEEE Transactions on Computational Social Systems, 2022, 9(5): 1563-1573.

[64] Qin R, Ding W W, Li J J, et al. Web3-based decentralized autonomous organizations and operations: architectures, models, and mechanisms[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2023, 53(4): 2073-2082.

[65] 王飞跃, 缪青海. 平行科学: 大模型时代 AI4S 的前沿技术与框架体系[J]. 学术前沿, 2024(14): 64-79.
Wang F Y, Miao Q H. Parallel science: cutting-edge technologies and framework system of AI4S in the era of large models[J]. Frontiers, 2024(14): 64-79.

[66] 王飞跃, 王雨桐. 数字科学家与平行科学: AI4S 和 S4AI 的本源与目标[J]. 中国科学院院刊, 2024, 39(1): 27-33.
Wang F Y, Wang Y T. Digital scientists and parallel sciences: the origin and goal of AI for Science and Science for AI[J]. Bulletin of Chinese Academy of Sciences, 2024, 39(1): 27-33.

[作者简介]



宫晓燕 (1976-), 女, 博士, 中国科学院自动化研究所多模态人工智能系统全国重点实验室高级工程师, 主要研究方向为复杂系统管理与控制、交通流预测、人工智能。



戴星原 (1993-), 男, 博士, 中国科学院自动化研究所多模态人工智能系统全国重点实验室副研究员, 主要研究方向为平行交通、平行驾驶、强化学习。



李芮霖 (2002-), 女, 山东交通学院硕士生, 主要研究方向为交通信号管控及交通管理。



吕宜生 (1983-), 男, 博士, 中国科学院自动化研究所多模态人工智能系统全国重点实验室研究员, 澳门科技大学创新工程学院教授, 主要研究方向为智能交通、无人驾驶、人工智能。