

引用格式: 罗斌, 刘文豪, 吴进, 等. 从地理信息系统到地理智能体[J]. 地球信息科学学报, 2025, 27(1): 83-99. [Luo B, Liu W H, Wu J, et al. From geographic information system to geographic intelligent agent[J]. Journal of Geo-information Science, 2025, 27(1): 83-99.] DOI: 10.12082/dqxxkx.2025.240658; CSTR:32074.14.dqxxkx.2025.240658

从地理信息系统到地理智能体

罗斌^{1,2}, 刘文豪^{1*}, 吴进¹, 韩嘉福^{1,2}, 吴文周¹, 李洪省^{1,2}

1. 中国科学院地理科学与资源研究所 地理信息科学与技术全国重点实验室, 北京 100101; 2. 北京星球时空科技有限公司, 北京 100083

From Geographic Information System to Geographic Intelligent Agent

LUO Bin^{1,2}, LIU Wenhao^{1*}, WU Jin¹, HAN Jiafu^{1,2}, WU Wenzhou¹, LI Hongsheng^{1,2}

1. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; 2. Beijing StarEarth Technology Co., Ltd., Beijing 100083, China

Abstract: [Objectives] The geographic system is an integrated framework encompassing natural and human phenomena and their interrelationships on the Earth's surface. While Geographic Information Systems (GIS) can digitally process these geographic elements, they face challenges in addressing rapidly changing geographic contexts with complex 3D structures. This is primarily due to the lack of bi-directional interactions between physical and informational spaces, as well as their reliance on predefined rules and historical data. In this paper, we propose the concept of a "Geographic Intelligent Agent" as an advanced form of GIS, which integrates embodied intelligence, self-supervised learning, and multimodal language modeling to improve environmental perception, spatial understanding, and autonomous decision-making. **[Methods]** The architecture of the geographic intelligent agent consists of three core components: multimodal perception, an intelligent hub, and an action manipulation module. These components collectively acquire comprehensive environmental information through sensor networks, perform complex situation reasoning using knowledge graphs and generative models, and enable real-time control and multilevel planning of the physical environment. To adapt to differences between virtual and real environments, the geographic intelligent agent is tested using the earth simulator and a test field platform, equipping it with stronger autonomous capabilities in complex and dynamic geographic contexts. **[Results]** This paper also demonstrates the implementation of geographic intelligent agent in spatial intelligence applications using the virtual digital human "EarthSage" as an example. **[Conclusions]** As a prototype of the geographic intelligent agent, "EarthSage" integrates modules such as the spatiotemporal Knowledge Graph (GeoKG) and a Cognitive Map Generation Model (GeoGPT), assisting users in obtaining intelligent spatial decision-making support in fields such as emergency management, urban planning, and ecological monitoring. This work exemplifies the transformation of GIS from a traditional information processing tool to an autonomous spatial intelligent system, marking a significant advancement in the field.

收稿日期: 2024-11-27; 修回日期: 2025-01-09.

基金项目: 国家重点研发计划项目(2021YFB3900901). [**Foundation item:** National Key Research and Development Program of China, No.2021YFB3900901.]

作者简介: 罗斌(1982—), 男, 湖南邵阳人, 硕士, 高级工程师, 系统架构师, 主要从事空间智能与认知地图、全息地球平台技术研究。E-mail: luob@igsrr.ac.cn

*通讯作者: 刘文豪(2000—), 男, 陕西西安人, 硕士生, 主要从事人工智能、遥感图像智能解译研究。
E-mail: liuwenhao22@mailsucas.ac.cn

Key words: intelligent geographic system; geographic intelligent agent; embodied intelligence; self-supervised learning; multimodal perception; knowledge graph; large model; spatial intelligence

***Corresponding author:** LIU Wenhao, E-mail: liuwenhao22@mails.ucas.ac.cn

摘要:【目的】地理系统是涵盖地球表层自然与人文现象及其相互关系的综合系统,而现有地理信息系统(GIS)虽能数字化处理这些地理要素,但其局限性在于缺乏物理与信息空间的双向交互,并且其模型通常依赖于预设规则和历史数据,难以应对快速变化和三维结构复杂的地理情境。为此,本文提出了“地理智能体”,作为地理信息系统的进阶形式,融合了具身智能、自监督学习和多模态语言模型,旨在提升环境感知、空间理解和自主决策能力。【方法】本文设计的地理智能体架构包含多模态感知、智能中枢和行动操控模块,分别通过传感器网络获取全方位环境信息、利用知识图谱和生成模型进行复杂情境推理,并最终实现对物理环境的实时调控和多层次规划。此外,地理智能体将通过地球模拟器和试验场平台测试,以适应虚拟和真实环境的差异,从而在复杂、动态地理情境中具备更强的自主应对能力。【结果】本文以虚拟数字人“地球通”为例,初步展示地理智能体在空间智能化应用中的具体实现。【结论】“地球通”作为地理智能体的原型机,集成了时空知识图谱(GeoKG)和认知地图生成大模型(GeoGPT)等模块,能够辅助用户在应急管理、城市规划和生态监测等领域中快速获取智能化的空间决策支持,充分体现了地理信息系统从信息处理工具向自主空间智能体的发展演化。

关键词: 智能地理系统; 地理智能体; 具身智能; 自监督学习; 多模态感知; 知识图谱; 大模型; 空间智能

1 引言

地理系统是由自然和人文要素构成的复杂有机体^[1-2],广泛分布于地球表层,具备生成、维持和转化物质与能量形态的功能。地理信息系统(GIS)作为其数字化表达工具,虽能对地理系统中的要素进行数字化处理,却只能将这些要素单向映射至信息空间,难以实现物理空间与信息空间之间的双向互动。随着数字孪生与人工智能技术的发展,GIS通过引入物联网技术,也试图构建物理世界与信息世界之间的双向映射体系,但在面对复杂、动态的地理环境时仍然存在显著局限^[3]。首先,受制于AI技术的发展,当前GIS依赖预先训练的模型,而这些模型通常基于历史数据和既定规则,缺乏快速学习和自适应能力,这在处理地理环境的连续变化时表现出不足,尤其在应对快速出现的不可预测事件时,系统响应滞后。其次,GIS的常用模型多基于二维或像素级别的预测,难以有效理解三维空间结构的复杂性。这种局限性阻碍了系统在高维空间中实现智能操作,难以提供更高级别的空间智能功能,如对全球跨尺度、复杂地形的精准理解和操作。最后,尽管GIS能够处理大量实时数据,但其学习机制主要依赖于外部监督,缺乏通过自监督学习机制来自主适应环境变化的能力。因此,系统在面对新环境或未曾遇到的情况时,难以进行有效的响应和调整。

智能化的GIS经历了早期GIS与自动化(1960—1980年)、GIS与机器学习的初步结合(1990—2010

年)、深度学习与大数据驱动的智能GIS(2010—2020年)3个主要发展阶段,如图1所示。随着智能地理系统的提出,物理空间与信息空间的双向交互、自主智能决策等问题得到了解决^[4-5]。智能地理系统通过双向映射,实现了物理世界与信息世界的交互,其核心框架由地理传感网、地理智能网和地理控制网三部分组成。地理传感网负责将物理世界实时映射到信息空间,依靠传感器网络收集地理环境的动态和静态数据,确保物理世界在信息系统中的精准呈现。地理智能网作为系统的核心处理单元,运用人工智能和数据分析技术,对传感数据进行处理、分析与判断,生成预测和决策建议^[6]。地理控制网则负责将地理智能网的决策反馈至物理世界,通过物联网技术对物理设备进行自动化控制,从而实现对现实地理环境的直接影响和调控。

当前,人工智能领域,具身智能^[7]、自监督学习^[8]和多模态大语言模型^[9]等技术快速发展,为GIS的智能化发展提供了新的路径,促进其从信息处理工具向自主空间智能系统转变^[10-11]。本文将这种具有空间智能的GIS系统称为“地理智能体”,旨在通过结合具身智能、自监督学习和大模型等先进技术,突破传统GIS的局限,赋予系统更强的自主感知与空间智能操作能力。其中,具身智能作为一种结合“身体”和“大脑”的智能技术,不仅强调通过传感器和执行器与环境进行直接交互,还借助人工智能算法提升系统对复杂环境的适应能力。自监督学习则是具身智能实现环境感知和适应性学习的关键技术之一,通过从未标注数据中提取模式,提高了

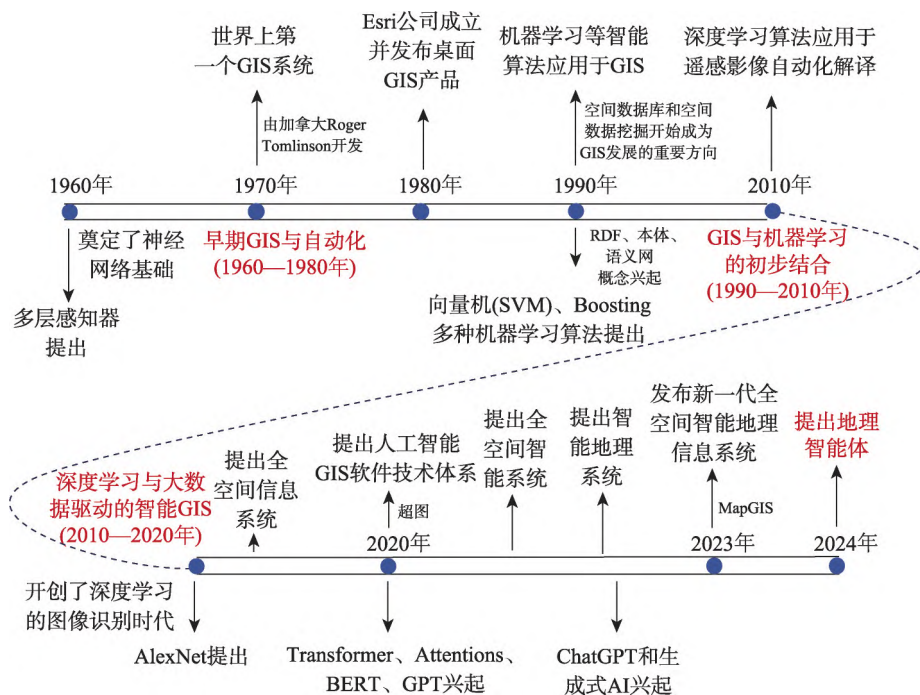


图1 智能GIS发展历程

Fig. 1 The development history of intelligent GIS

系统在复杂情境中的学习和调整能力^[12]。结合具身智能与自监督学习^[7],地理智能体能够自主感知三维地理环境并自动化空间数据生成,逐步适应其动态变化,具备灵活应对复杂事件的能力,尤其在突发事件中表现出高效响应的优势^[13]。此外,地理智能体不同于传统的像素级别预测,能够实现对三维地理结构的深度理解与本源分析,提供可信自主决策与智能代理服务^[14]。因此,地理智能体的提出将标志着地理信息系统从静态的信息处理工具转变为具备动态决策、与现实交互和实时适应能力的智能系统,为实现空间智能迈出关键的一步。

2 地理智能体定义与框架

2.1 定义与框架

地理智能体作为一种结合人工智能新技术的智能地理系统,不仅继承了智能地理系统的双向映射功能,还在此基础上进一步引入了具身智能的概念。通过传感器网络 and 智能控制机制,地理智能体能够实时捕捉物理地理系统的变化,在数字空间中精准映射,并将数字空间的决策高效反馈至物理世界。

地理智能体是智能地理系统的一种新的智能化形态,融合了具身智能、自监督学习和多模态大语言模型等技术,其核心特征在于自主性和自适应

性,不仅能实时感知三维地理环境,还能进行复杂情境推理和自主决策,与传统GIS系统主要依赖预定义规则和历史数据的方式不同,地理智能体通过知识图谱和生成模型等技术,具备更强的学习能力和适应能力,能够动态响应复杂和快速变化的地理环境,这种升级使地理智能体能够更灵活地应对突发事件,特别是在应急管理和生态监测等场景中展现出更高的效能。地理智能体的架构如图2所示,其智能主体主要由多模感知、智能中枢和行动操控3个部分组成。

多模感知模块通过整合天基、空基、地基、岸基、水面、水下及海底传感器网络(如星载遥感、无人机影像、地面摄像头、激光雷达和原位传感器)以及互联网数据源,实现对环境的全方位、多维度实时感知和数据采集。与传统的传感网相比,多模感知不仅局限于单一维度的数据获取,还能感知包括视觉、音频、语音、触觉、温湿度、振动、气体成分、激光扫描等多种数据类型。例如,视觉感知可以通过遥感影像、无人机视频和地面摄像头捕获高分辨率图像;声学感知通过捕捉地震波、风暴噪声等信号,帮助识别自然灾害的早期迹象;触觉感知则通过气象站、温湿度传感器等设备实时监测关键地理要素的动态变化。这些扩展感知能力显著提升了系统对复杂地理现象的识别和分析能力,确保多模感知模

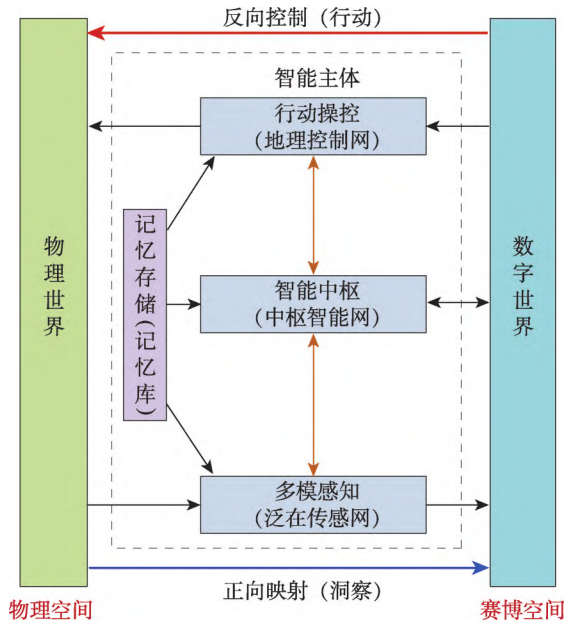


图2 地理智能体的架构

Fig. 2 Architecture of geographic intelligent agent

块不仅仅是传统数据获取手段的叠加,而是实现了对环境的深层次理解与动态响应,为智能中枢提供更加全面和多样化的时空数据支持。

智能中枢模块作为核心决策引擎,由知识图谱模型和生成式大模型组成,负责对多模感知模块收集的数据进行分析、语义推理、情境模拟与情景预测。知识图谱模型专注于空间关系的精确表达、逻辑规则和路径分析,能够细致整合来自各个传感器网络的数据,为复杂地理任务提供高精度的逻辑推理和优化方案。生成式大模型则聚焦于全局空间认知与推理创造,通过对多模态数据进行深度学习与情境分析,基于复杂地标和情境信息提供全局性空间决策支持。二者协同运作,结合大数据挖掘、深度学习和自监督学习等技术,实现对多源复杂地理数据的高效处理,并生成精准的智能决策指令。例如,在灾害应急响应中,智能中枢通过知识图谱提供灾区的地理关系网络,结合生成式大模型对动态数据的全局分析,生成资源调配和路径优化的全局规划方案。

行动操控模块是地理智能体的核心执行层,通过智能化控制网络将智能中枢生成的决策快速反馈至物理环境,并完成复杂任务的动态执行和调控。行动操控模块结合具身智能、自监督学习和多模态大模型技术,具备实时响应、自适应调整和高效率任务执行的能力。通过智能化调度管理物联网设备(如无人机、地面机器人和传感器网络),该模

块能够实现任务分解与动态调整,并根据实时反馈优化执行策略。例如,在洪水灾害场景中,行动操控模块可根据传感器实时监测的水流变化,自主调整监测设备的部署位置,并生成动态优化的救援路径。通过具身智能技术,行动操控模块能够实现环境与行为的紧密耦合,使系统能够适应复杂地理环境的动态变化。此外,自监督学习技术将进一步增强了模块对未标注数据的自适应能力,通过持续学习优化行动策略,提高任务执行效率和精确性。

因此,多模感知、智能中枢和行动操控三大模块通过密切协同构建了地理智能体的核心功能闭环。多模感知提供全方位数据支持,智能中枢生成逻辑推理与全局规划,行动操控模块完成任务执行与反馈优化。三者的协同作用使地理智能体能够在复杂动态环境中完成多样化任务,展现出卓越的空间感知、自主决策和智能执行能力,为智慧城市、生态保护、灾害管理等多领域的应用提供了技术支撑与理论指导。

2.2 多模感知模块

多模感知模块是地理智能体实现物理与信息空间深度交互的核心模块,通过整合天基、空基、地基、岸基、水面和水下等多源传感网络,实时感知地理环境的动态和静态特征。相比传统的地理传感网,地理智能体的多模感知模块不仅在感知范围上覆盖更广,还能够采集多样化的感知数据,包括视觉、声学、触觉、激光以及互联网动态数据。这种多模态感知手段使系统能够在复杂的地理环境中,主动从多角度感知地理现象,为后续的智能决策提供高质量的时空数据支持。

多模感知模块的特点在于其强大的多模态数据融合能力。通过深度学习与数据挖掘技术,多模感知能够将来自不同传感器的数据整合为统一的时空信息。比如,在洪水监测中,视觉感知提供洪水覆盖范围的遥感图像,触觉传感器记录水流压力变化,声学传感器捕捉水流声波信号,而互联网数据提供实时的公众反馈信息。通过这些数据的融合分析,系统可以更全面地描述洪水的发生过程及其影响范围,从而提升对复杂灾害的认知能力。

此外,多模感知模块具有显著的协同感知和实时响应能力。针对同一地理现象或过程,系统能够通过不同模态的数据源实现多角度协同感知。例如,在森林火灾监测中,系统可利用高分辨率遥感影像捕捉火灾范围,结合温湿度传感器监测气候变

化,再通过烟雾检测和社交媒体动态反馈评估火灾传播速度。这种协同感知能力不仅提高了地理智能体对突发事件的响应速度,还增强了其在动态环境中的适应性和稳定性。

多模感知与智能中枢之间形成了明确的分工。多模感知模块聚焦于对外界环境的主动感知和实时数据采集,是地理智能体的“感官系统”。智能中枢则专注于对这些数据进行语义化分析、复杂情境推理和智能决策生成,是系统的“思维中枢”。两者协同合作,通过感知与决策的闭环体系,实现了对复杂动态地理环境的精准感知和高效响应。

2.3 智能中枢模块

智能中枢模块是地理智能体的“中枢智能网(核心神经中枢)”,负责接收、处理和分析地理环境的多维数据,并生成智能化决策。该系统借鉴人类大脑在空间认知中的功能分化^[15],采用知识图谱模型和生成式大模型作为智能中枢的“左脑”和“右脑”,其中,“左脑”主要负责逻辑与分析,“右脑”主要负责推理、创造和空间感知。二者相互协同,以实现地理空间的深层理解、合理规划和有效执行。

生成式大模型作为“右脑”,擅长从非结构化数据(如遥感图像、文本描述和实时动态数据等多模态数据)中提取语义信息,并将这些信息与知识图谱中的已有知识进行对接。它可以通过调用知识图谱中已有的地学信息与规则,从而在复杂情境中减少对外部监督的依赖,快速生成语义一致的推理结果。此外,生成式大模型通过分析非结构化数据,可以生成新知识和更新现有知识,以动态补充知识图谱的内容,使地理智能体在复杂多变的空间环境中更具灵活性和适应性。

知识图谱^[16-18]模型则扮演“左脑”角色,注重逻辑和细节导向的空间处理,用于构建与管理结构化的空间实体及其关系网络,例如道路网络、建筑分布、地理实体间的距离和连接性等。借助这种结构化的形式,知识图谱使生成式大模型的推理能够遵循明确的规则和上下文逻辑,提升地理智能体决策的准确性与可靠性。此外,借助图嵌入技术,知识图谱可为大模型提供优化后的数据表示,确保非结构化数据与结构化知识的高效融合。

两者协同的典型机制体现在任务分解与执行闭环中,它们通过数据流与任务目标的协同紧密结合。知识图谱先将复杂任务分解为逻辑清晰的子任务,为生成式大模型提供高质量的结构化语义信

息输入,确保其推理基于准确的空间数据;生成式大模型则根据环境变化生成动态决策建议,并通过知识图谱校验其逻辑一致性与可行性,进一步实现动态环境中的复杂情境推理与生成规划。这种协同机制将显著提升地理智能体在复杂环境中的任务执行能力,为其在动态地理环境下的智能决策提供了可靠保障。

2.4 行动操控模块

行动操控模块是地理智能体的核心执行模块,负责将智能中枢生成的决策快速反馈至物理环境,并完成复杂任务的执行与动态调控。通过智能化控制网络,行动操控系统可以实现虚拟空间与物理空间之间的深度联动,支持实时响应、自适应调整 and 高效任务执行。作为地理智能体架构中的关键环节,行动操控系统结合了物联网技术、具身智能和自监督学习等前沿技术,使其在多样化的地理环境中表现出卓越的适应性和灵活性。

行动操控模块的核心架构包括数据输入与指令接收、执行与控制以及反馈与优化三大子模块。数据输入模块负责接收智能中枢生成的多模态决策信息(如路径规划建议、资源分配策略等),并整合多模感知模块提供的实时环境数据(如地理变化、设备状态和用户反馈)。这些多源数据为系统的任务执行提供了全面的支持。执行与控制模块通过智能化调度和管理物联网设备,实现任务分解与动态调整。例如,在灾害应急响应中,行动操控系统可以指挥无人机与地面机器人协同作业,完成受灾区域的评估与救援。反馈与优化模块通过实时监控任务执行结果,将反馈数据传递至智能中枢,用于更新决策依据,并通过自监督学习技术持续优化执行模式。

行动操控模块的实现依赖于以下关键技术。首先,具身智能技术赋予系统在动态环境中通过传感器与执行器实现实时交互的能力。通过模拟生物体感知与行动的耦合机制,具身智能技术使系统能够主动感知周围环境的动态变化,并实时调整执行策略。其次,自监督学习技术使行动操控系统能够从未标注数据中自主提取有效模式,并动态优化任务执行策略。通过利用环境中大量未标注的传感器数据,自监督学习技术不仅降低了对外部监督的依赖,还增强了系统的自适应性。此外,多模态大模型技术为行动操控提供了对复杂环境的多维度理解能力。通过整合视觉、音频、触觉等多模态

数据,系统能够生成动态优化方案,并实时调整任务执行过程。

通过整合具身智能、自监督学习和多模态大模型技术,行动操控模块实现了对复杂地理环境的灵活适应和高效响应。其与智能中枢和多模感知模块的协同机制,将显著增强地理智能体在多场景、多任务中的执行效率与决策质量,可为动态地理环境下的智能化应用提供关键技术支撑。

3 地理智能体构建思路

地球科学的研究范式经历了从基于物理的模型向数据驱动模型的转变^[19]。传统的基于物理的模型建立在物理和数值框架之上^[20],通过精确重建潜在的物理过程为地理现象或规律提供了可靠的解释能力。然而,随着大数据技术的兴起,当代研究范式逐渐转向依赖数据驱动方法,特别是利用机器学习和深度学习技术。数据驱动模型能够从海量地球科学数据中提取出重要见解,无需对复杂理论进行详尽建模。尽管机器学习在智能地理系统构建方面展现出巨大的潜力,但数据稀缺、计算资源需求高、数据隐私问题,以及人工智能模型的“黑箱”属性等挑战,仍然限制了其广泛应用。因此,将基于物理的模型与数据驱动模型相结合,形成混合模型,是未来人工智能与智能地理系统深度融合构建地理智能体的关键路径。

如图3所示,地理智能体的具身形式可采用虚拟数字人或其他适配地理学需求的形式。虚拟数字人作为一种人机交互的表现形式,能够通过自然语言处理与多模态感知技术,为用户提供便捷的空间知识服务。然而,仅采用虚拟数字人的形式可能会使地理智能体与地理学研究的实际需求脱节。因此,在具体应用场景中,地理智能体的具身形式应根据任务需求灵活调整。例如,在生态监测场景中,地理智能体可以呈现为与环境传感器网络深度集成的控制单元;在城市规划中,可以表现为虚拟城市助手,直接通过多维可视化交互支持复杂空间规划与决策。这种多样化的具身形式设计,使地理智能体更贴近地理学实际问题,同时避免形式单一化导致的局限性。

3.1 地理智能体设计原则及总框架

3.1.1 设计原则

地理智能体的设计需遵循自主性、适应性和扩展性三大关键原则,但这些原则需要结合地理学的具体需求加以实现,以更好地服务于复杂的地理环境分析与管理。与传统GIS的静态信息处理能力相比,地理智能体通过自主性增强了应对动态地理环境的能力,包括实时感知、情境推理和决策执行等。然而,自主性并非其唯一或核心区别,而是与适应性和扩展性共同形成了对复杂地理问题的系统化解决方案。

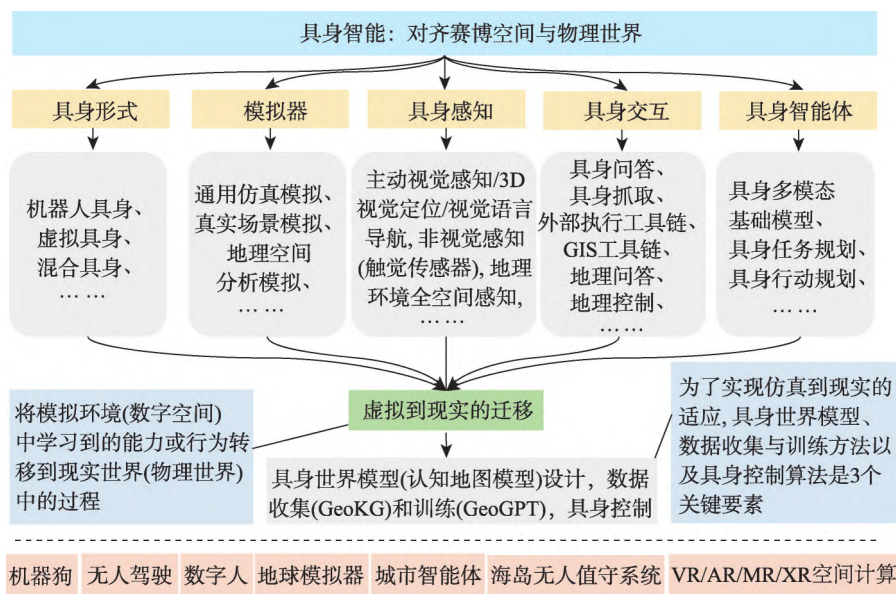


图3 智能地理系统具身智能体系组成

Fig. 3 The components of the embodied intelligence in intelligent geographic systems

自主性要求地理智能体在无需大量人工干预的情况下,能够独立完成复杂地理任务,这不仅需要依赖数据驱动模型对时空数据的模式识别能力,还需结合基于物理的模型对地理过程的内在机理进行解析。例如,在突发环境事件中,地理智能体可以自主分析多源数据,生成优化后的应急方案,而无需依赖预定义规则。适应性要求地理智能体能迅速响应新的环境变化,数据驱动模型(尤其是机器学习和深度学习)能够从大量地理数据中提取模式,增强地理智能体在不确定情境中的应对能力。物理模型的引入将为数据稀缺场景中的预测提供了稳定的基础,自监督学习则进一步提升了地理智能体在数据稀缺情况下的学习能力,使其结构能够在缺乏标注数据的情境中不断优化,自主更新以适应环境变化。扩展性则确保了地理智能体能够通过模块化设计适应不同的地理学研究场景,支持从城市规划到灾害管理等多领域应用的适配性。

3.1.2 总框架

在实现这些原则的过程中,关键在于如何构建地理智能体系统内部用于认知世界、表示和理解其所处环境的抽象模型,即人工智能世界模型,帮助地理智能体通过感知、学习、推理和预测来理解环境的状态及其动态变化,从而能够进行自主决策、规划和互动。世界模型不仅仅是对环境的静态描述,还包括对环境演变的动态预测、不同因素之间的相互关系以及行动的潜在后果。认知地图^[21-23]作为一种借鉴人脑结构设计的动态心智表征形式,为地理智能体空间智能世界模型的构建提供了独特的解决方案。最初,认知地图用于解释动物具备的空间导航能力,其核心在于构建对空间与环境属性的动态理解。随着认知科学和人工智能的快速发展,认知地图的应用逐步扩展到更广泛的领域,包括空间知识的组织、环境信息的推理与决策支持等。因此在地理智能体的三大原则实现中,认知地图模型能够为组织空间知识与环境属性提供高效的结构化支持。其动态表征能力不仅能够帮助地理智能体更好地理解复杂环境,还能通过推理和导航能力增强其在复杂情境中的决策水平,从而全面提升适应性和扩展性。

如图4所示,在设计中,认知地图由知识图谱模型和多模态大模型协同构建。知识图谱模型利用基于物理的模型,在认知地图中构建结构化空间关系的精准表达和逻辑推理,使智能体在地理

决策中能够使用物理过程的显性知识,进行路径规划、资源分配等任务,为智能体提供逻辑严密的分析支持,并根据真实环境的地理约束(如地形特征、道路通行性)对大模型生成的解决方案进行逻辑校验与优化,从而避免错误决策。多模态大模型则依托数据驱动方法,通过深度学习处理非结构化的多模态数据,从中提取语义信息以构建高维情景向量,为认知地图提供动态环境预测和全局语义补充。通过多模态大模型,认知地图能够准确捕捉多模态数据间的隐含关联,为复杂情境中的情景分析提供语义一致性支持。因此通过多模态大模型与知识图谱的集成,认知地图在面对复杂、多变的地理环境时,将具备推理能力和丰富的情景理解与预测能力,这种设计不仅可以提升地理智能体对多模态信息的整合与处理能力,还会显著增强其在动态环境中的决策准确性和响应效率^[24]。

此外,物联网与智能控制技术为地理智能体提供了对物理设备的实时控制和管理能力。物联网技术使智能体能够高效执行认知地图中的决策指令,而智能控制技术则提高了行动操控的灵活性,使智能体在不同情境中能够自动优化执行策略。通过这一协同框架,认知地图充当复杂地理空间的类人化认知模型,通过多层次的数据整合和语义推理,使地理智能体能够在环境中自主感知、动态适应并灵活应对多种地理情境。

3.2 地理智能体研究支撑环境

3.2.1 地球模拟器

地球模拟器作为一种高效、经济且可控的仿真平台,可为地理智能体提供一个模拟现实场景的虚拟环境,用于支持地理智能体的训练、决策和实验,使地理智能体能够在可控、虚拟的环境中进行多样化的测试与训练,从而大幅降低实际操作中的风险与成本。地球模拟器可以在受控环境中复现自然灾害或极端地理现象,保障了地理智能体的安全性和稳健性。此外,地球模拟器具备快速构建地理模型和原型的能力,加速了环境模型的开发与优化。凭借其开放性,地球模拟器还为地学研究社区提供了共享的平台,促进跨学科的协同创新与数据共享,为不同领域的研究提供了实证支持。地球模拟器的架构如图5所示,按照“观测、假设、预测”运行框架,地球模拟器整合数据网络、数字孪生时空大

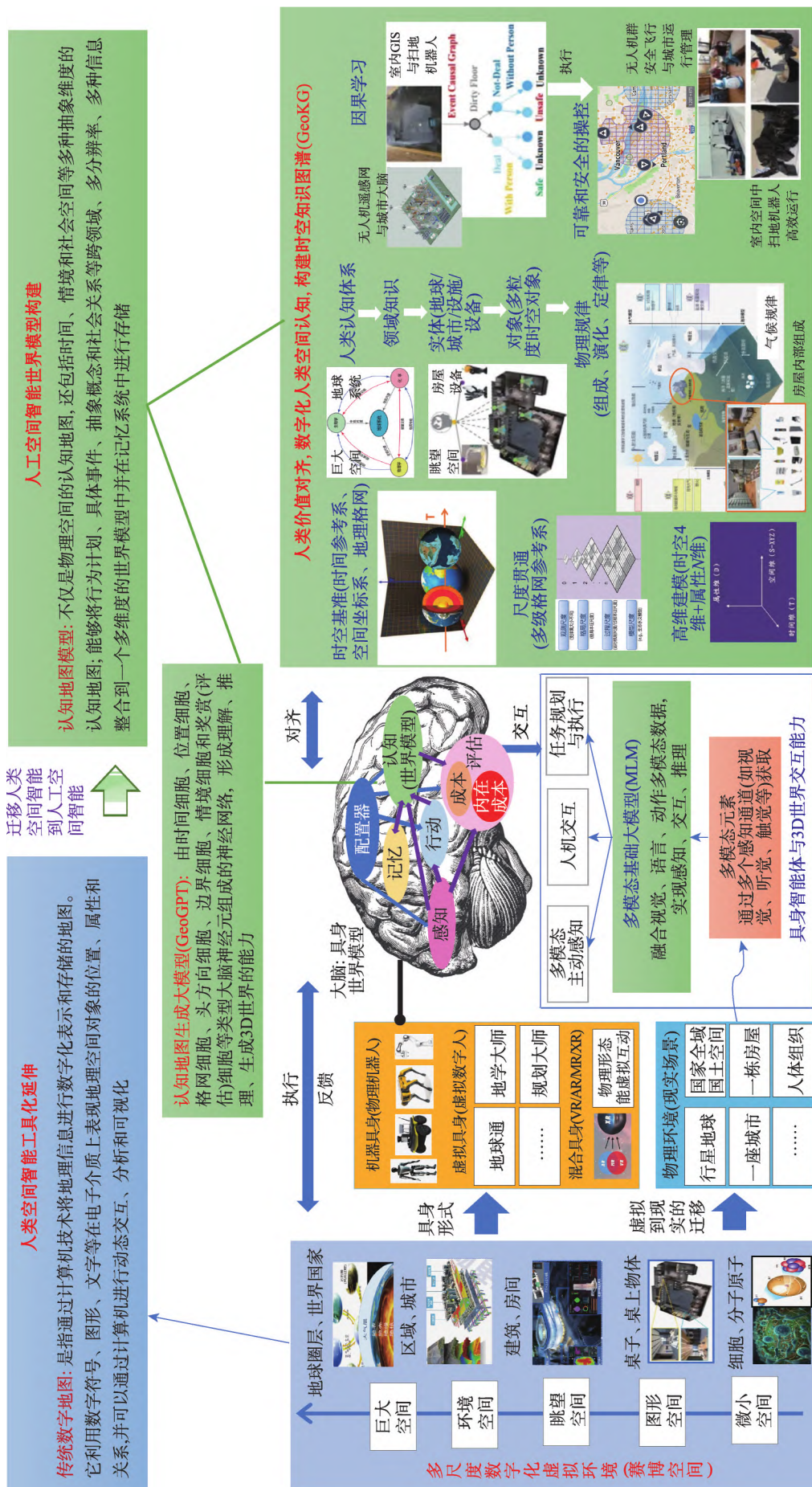


图 4 基于认知地图模型的地理智能体总框架
Fig. 4 A general framework for geographic intelligent agent based on cognitive map models

数据平台、地理空间“数据到智能”演化全生命周期建模、监测评估预测分析模拟工作台等技术构建,不仅能够支持地球圈层耦合大尺度分析模拟,而且能够支持城市运行、街区导航等精细尺度分析模拟。

大尺度地球系统模拟主要通过集成全球多源观测数据,包括天基遥感、空基无人机影像和地基传感器网络,地球模拟器实现了对地球系统的全方位感知。其观测范围覆盖了全球气候、大气环流、海洋动力学等宏观环境变化,同时深入到更精细的生态系统动态和城市扩张等局部现象。图5中的“地理空间建模”展示了这一数据整合过程,通过实体对象表达模型对所有观测数据进行融合,并按照维度化集成建模构建地球数据立方体,实现数据的对齐;进而地球模拟器通过知识图谱引擎与生成式大模型技术,将多源观测数据转化为知识化表达深度神经网络和知识图网络,实现了从环境感知到认知建模的跨学科整合。这种数据耦合与整合能力,使地球模拟器能够生成统一的环境模型,为全球气候预测、海洋变化模拟和生态趋势分析提供有力支撑^[25]。在这样的数据支持下,地球模拟器可以运行全球气候模型(如ICON模型),生成气候变化的中长期预测,帮助决策者实时掌握气候变化动态,为应对台风、洪水等极端天气事件提供可靠的情景输入。通过整合观

测数据和复杂自然系统的模拟,地球模拟器在宏观尺度上帮助地理智能体构建对地球系统的全面理解与预测能力,支持其在地球系统模拟中的多种应用。

在精细尺度上,地球模拟器通过高保真的虚拟测试环境,支持自动驾驶技术的开发与测试。地球模拟器的“预测与假设”模块可以模拟复杂的地球系统,同时在局部场景中为自动驾驶智能体提供逼真的测试条件。在自动驾驶模拟中,地球模拟器通过构建虚拟的城市道路、乡村公路等多样化交通场景,为自动驾驶系统提供一个全面的虚拟测试平台,确保其在不同天气和光照条件下、特别是灾害天气条件下的感知、决策和控制能力得到充分验证。此外,模拟器能够模拟自动驾驶传感器(如摄像头、激光雷达、毫米波雷达)的虚拟输出,提供高度精确的虚拟数据,以帮助验证自动驾驶算法的鲁棒性和准确性。更重要的是,模拟器支持多智能体协同仿真,能够模拟复杂的交通流和动态交互,使自动驾驶智能体能够在多种路况下应对行人、车辆和骑行者等多类交通参与者的互动情境。通过这种精细尺度的模拟,地球模拟器不仅加速了自动驾驶技术的开发进程,还为其在实际应用中的安全性和可靠性提供了坚实保障。

3.2.2 试验场

地球模拟器虽然为地理智能体提供了高效、经

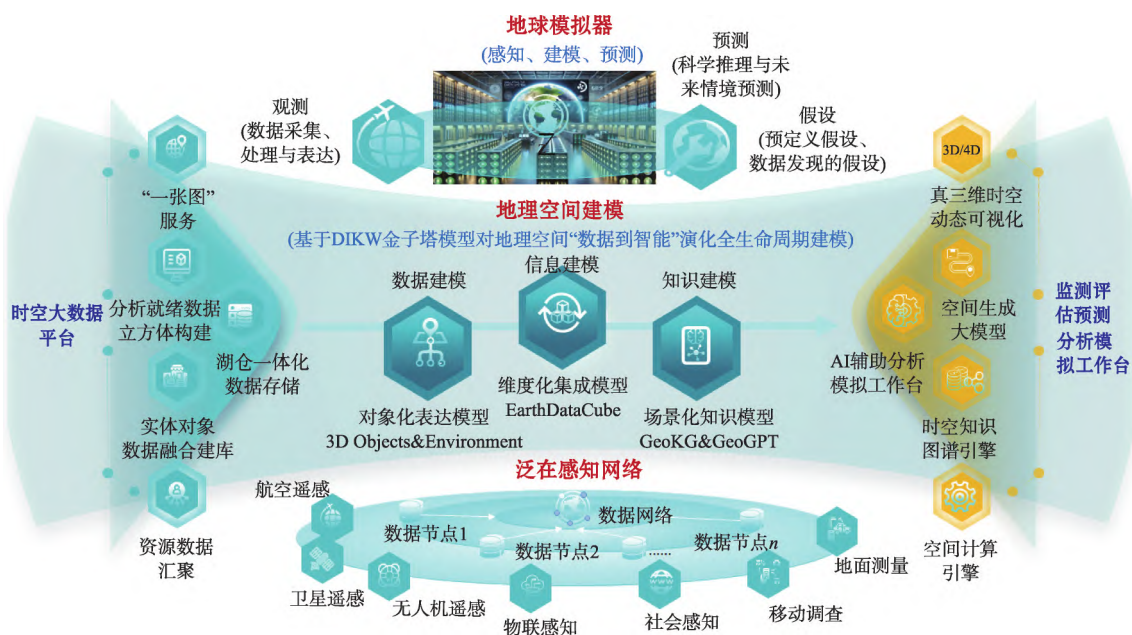


图5 地球模拟器架构

Fig. 5 Architecture of the earth simulator

济且可控的虚拟测试环境,但在仿真过程中难以完全再现现实世界中的突发自然扰动、意外的人类活动干扰以及气候和地形的微小变化等因素,而这些因素对地理智能体的实际应用表现至关重要。而试验场作为地理智能体的重要测试平台,相较于地球模拟器,可以为地理智能体在真实地理环境中提供真实的验证条件和应用反馈。试验场通过在自然环境、人造环境及多种实际场景中测试,弥补了地球模拟器在精细尺度上的不稳定性,确保地理智能体在复杂多变的真实物理环境下具备稳健的表现和自主应变能力。

搭建试验场需要结合自然地理环境、人造环境以及多种实际场景的模拟。这不仅要求考虑地理环境的多样性,还需要综合不同地域的气候特征、地形变化以及人类活动干扰因素。因此,搭建试验场不仅依赖于政府场地的支持,还需要相关研究机构提供技术协助,确保平台能够全面模拟不同地理条件下的工作情境。在具体实现上,试验场将由多个功能区组成,包括自然地理试验区、城市试验区、农业与乡村试验区以及灾害应急试验区等。每个区域都配备传感器、环境监控系统以及数据采集设备,进行实时监测并将数据传输至中央控制系统进行分析。此外,试验场还将配置专门的测试设施,如高精度传感器、模拟气候系统和智能设备,以确保全面评估地理智能体在各种环境中的适应性和鲁棒性。

通过在试验场中运行,地理智能体可获得真实场景的反馈数据,能够在不同情境下不断调整和优化其算法与系统功能,补充了地球模拟器所无法覆盖的测试需求。这种真实环境的实践不仅显著提升了智能体的稳定性和可靠性,还为未来地理智能体在实际应用中的大规模部署奠定了更坚实的科学基础。

3.3 地理智能体构建关键技术

在地球模拟器与试验场的基础上,地理智能体的构建依赖一系列关键技术,除了地理空间专业领域人工智能技术突破之外,结合空间智能认知地图世界模型的具身智能技术最为关键,包括具身感知技术、具身交互技术、具身代理技术,以及从模拟到真实环境适应的具身控制技术。这四大技术既具有独立的实现路径,又通过协同作用共同构建了地理智能体在复杂地理环境中的感知、交互、决策和适应能力。

3.3.1 具身感知技术

具身感知技术是地理智能体构建的关键基础之一,其核心在于赋予智能体主动感知地理环境的

能力,并通过动态、多尺度的感知机制支持复杂地理任务的实现。这种技术不仅需要地对地理环境进行精细的空间理解,还需动态适应地理环境的变化,保证感知结果的实时性和可靠性。通过多模态传感器网络,地理智能体可以主动获取环境信息,突破传统被动数据接收模式的限制。天基、空基和地基多模态数据(如遥感影像、无人机数据、地面传感器信息及实时天气和交通数据)的融合,使地理智能体能够实现从宏观到微观的多尺度感知,提供实时的三维视觉信息。例如,Jayaraman等^[26]提出了一种强化学习方法,使智能体通过减少对未观察到的部分的不确定性,主动获取信息性视觉观察,完成全景场景和3D物体形状的重建。

此外,地理智能体还可以借助深度学习和自监督学习技术,从未标注或部分标注的地理数据中提取复杂模式。与传统依赖大量标注数据的方法不同,自监督学习显著增强了智能体在地理数据稀缺条件下的适用性。例如,在地理灾害预警中,智能体通过从历史地震影像中学习场景变化模式,即使在面对未知区域或少量数据时,仍能准确感知滑坡、泥石流等灾害的潜在风险。Tatiya等^[27]提出的多阶段投影框架已被改进应用于此类场景,能够通过探索性交互传递隐性知识,使智能体更高效地感知地形动态特征。边缘计算与分布式计算技术可以进一步提升具身感知系统的效率,通过分散数据处理任务减少传输延迟,并加快响应速度。这种技术架构可以为智能体的实时感知、预测和决策奠定了基础。同时,多模态大模型的引入可以极大增强地理智能体处理复杂数据的能力,显著提升地理智能体在复杂地理场景中的实时感知和智能化水平。

3.3.2 具身交互技术

具身交互技术是地理智能体构建中支持其动态决策和实时响应能力的关键环节。这项技术通过构建地理智能体与物理和数字地理环境之间的双向交互机制,使地理智能体能够实时感知环境变化并做出响应,满足动态地理任务需求。地理智能体的具身交互不仅依赖多模态数据的综合分析,还强调交互过程中认知能力与行动规划的协同优化。例如,在城市交通管理中,地理智能体可以通过实时分析交通流量和信号灯数据,主动调整信号灯时长,实现道路网络的动态优化。Gordon等^[28]提出的分层交互式记忆网络就可用于此场景,该网络使负责任务选择的规划器和执行任务的低级控制器交

替运行,实现智能体与动态环境之间的交互。借助此类方法,地理智能体可通过多模态数据(如视觉、空间和语言信息)的融合处理,形成对物理与数字环境的全面理解。

具身交互的核心在于地理智能体的持续学习能力,即通过与环境的反复交互优化自身的行为策略,这需要基于实时的具身感知反馈与交互优化。在复杂地理场景中,地理智能体可以通过感知-决策-行动的闭环交互机制,逐步优化其交互策略。例如,在无人机辅助的地理数据采集任务中,智能体可以实时分析采集数据的分布情况,根据区域的地形特征和环境变化动态调整航线规划,以提高数据覆盖率和采集效率。而对于难度极高的深海矿产勘探或太空资源采集任务,配备具身交互技术的地理智能体可以通过机械臂抓取^[29]等方式,克服极端环境下依赖互联网和人力远程操作的局限性,确保在复杂地理场景中的长期可靠性与高效性。

3.3.3 具身代理技术

具身代理技术赋予地理智能体自主感知环境、制定决策并采取行动以实现特定目标的能力,是地理智能体构建中不可或缺的关键技术。具身代理的核心在于自主性和适应性的高度统一,使地理智能体能够在复杂多变的地理场景中独立完成任务,代表用户或系统高效运行。其能力不仅局限于物理和网络环境中的交互,更涵盖了自主决策、动态任务规划以及执行能力。

具身代理的关键在于感知、决策和行动的紧密耦合。其中,认知地图的动态更新可为地理智能体提供实时环境感知与任务规划的基础框架;强化学习技术通过试错与反馈机制,使智能体能够在迭代中不断优化策略,提升任务完成效率与环境适应性;多模态大模型的引入将进一步增强了地理智能体的环境理解能力,使其能够从动态地理数据中提取有用信息并灵活应对复杂场景。例如,PaLM-E^[30]开发了集成多种感知模态和自然语言处理能力的模型。它通过视觉识别周围环境、通过听觉理解指令并理解自身状态,以实现复杂的交互和操作。Translated LM^[31]可以将复杂任务分解为可管理的简单步骤,并使用其内部逻辑和知识系统设计解决方案,而无需额外训练。这类模型在洪水监测等多种地学场景中具有重要应用价值。例如,地理智能体可通过强化学习技术优化巡航路线,同时结合实时传感器数据精确判断潜在危险区域,实现自主监测和任务执行,而无需依赖人工操作。

3.3.4 模拟到真实适应的具身控制技术

模拟到真实适应是指地理智能体将在模拟器中学习到的能力或行为转移到真实世界场景(物理世界)的过程,旨在解决仿真环境中学习的算法和模型如何高效迁移至真实世界的问题。通过在虚拟环境中进行训练,地理智能体能够提前积累任务经验,并通过算法与策略的优化适配真实物理场景中的复杂环境变化。这一技术将为地理智能体在实际任务中的应用提供核心支撑,尤其是在资源受限、环境复杂的地理场景中。

模拟到真实适应技术的核心在于通过模拟器的构建和强化学习算法的优化,形成地理智能体在真实环境中的行为基础。地球模拟器作为技术实现的关键载体,通过整合多模态数据,可为地理智能体提供一个与真实世界高度接近的学习环境。在这一环境中,地理智能体能够通过认知地图的动态更新模拟真实地理场景的变化,从而实现对地理环境中复杂动态特征的学习与适应。而强化学习技术将引导地理智能体在虚拟环境中不断优化其行为模式,使得地理智能体能够从有限的模拟数据中提取高价值的知识结构,并通过不断优化的策略增强其应对复杂任务的能力。例如,Real2Sim2real^[32]通过在模拟器中训练强化学习模型,增强现实场景中的模仿学习能力。该方法先利用模拟环境中采用的强化学习训练强化策略,再将这些策略迁移至现实世界,用于解决数据稀缺问题,并实现高效的智能体模仿学习。此类具身控制技术将使地理智能体能够以低成本、高效率的方式将虚拟环境中的学习成果应用于实际地理场景,为地理智能体的全场景适应奠定更为坚实的技术基础。

4 案例研究:虚拟数字人“地球通”

考虑到构建地理智能体是一项巨大工程,为了验证本文提出的构建思路,采用虚拟数字人具身形式、对话机器人系统构建地理智能体的总体方案,基于全息地球^[33]平台(www.DeepEarth.cn),调用平台API和工具链简化实现智能体多模感知和行动操控模块;通过引入时空知识图谱(GeoKG)和认知地图生成大模型(GeoGPT),扩展通用智能代理(Agent)“感知、记忆、认知、行动”对于时空数据的处理能力^[34],从而构建一个能够提供时空知识服务的对话机器人“地球通 EarthSage”,实现框架如图6所示。

“地球通”通过对话机器人、虚拟数字人2种方

式进行人机交互,运行在PC桌面端和手机、平板等移动端,作为“认知世界的智能助手”,“地球通”具备更高的智能化和自主化优势。以建筑物提取任务为例,传统GIS方法经历了从使用ArcGIS、ENVI等软件内置的机器学习方法,逐步发展为采用深度学习模型实现高精度提取的演变。然而,这些方法通常只提供最终的提取功能,用户仍需自行完成数据获取与预处理、模型选择及结果后处理。尤其是

深度学习模型对输入数据格式的要求非常严格,这进一步增加了数据处理的复杂性和繁琐程度。与此相比,“地球通”实现了一站式处理流程,极大提升了用户体验和效率。图7显示了“地球通”与用户交互的一个实例。当用户需要提取某个区域的建筑物时,只需通过对话输入问题,“地球通”就可以通过全息地球平台主动跟踪获取互联网上相关位置的卫星遥感影像、自动提取该位置的建筑物,并

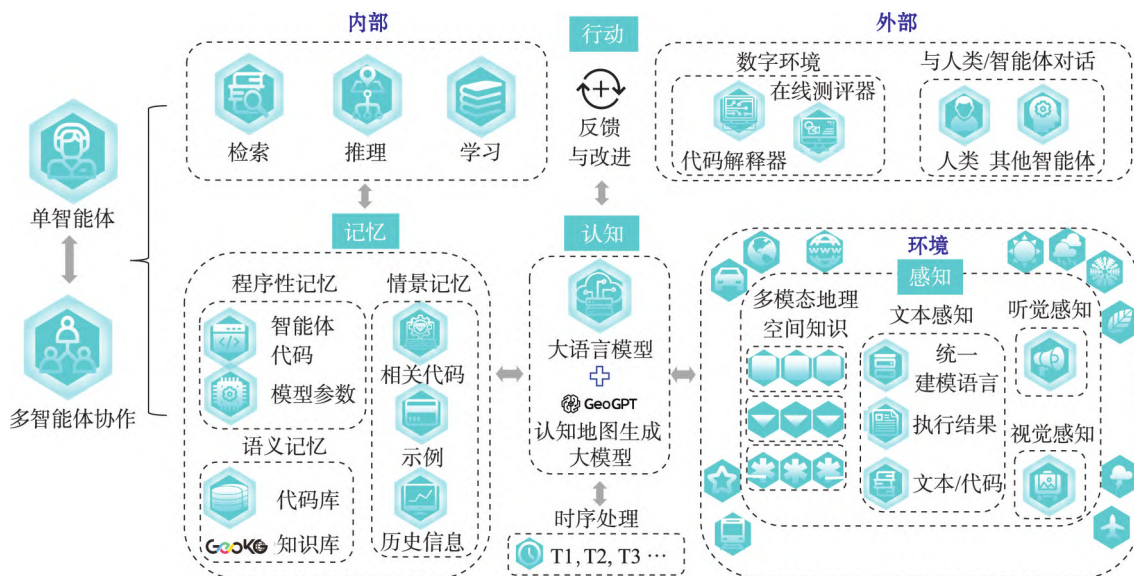


图6 智能代理Agent扩展GeoKG与GeoGPT的框架

Fig. 6 Framework of intelligent agent extension for GeoKG and GeoGPT

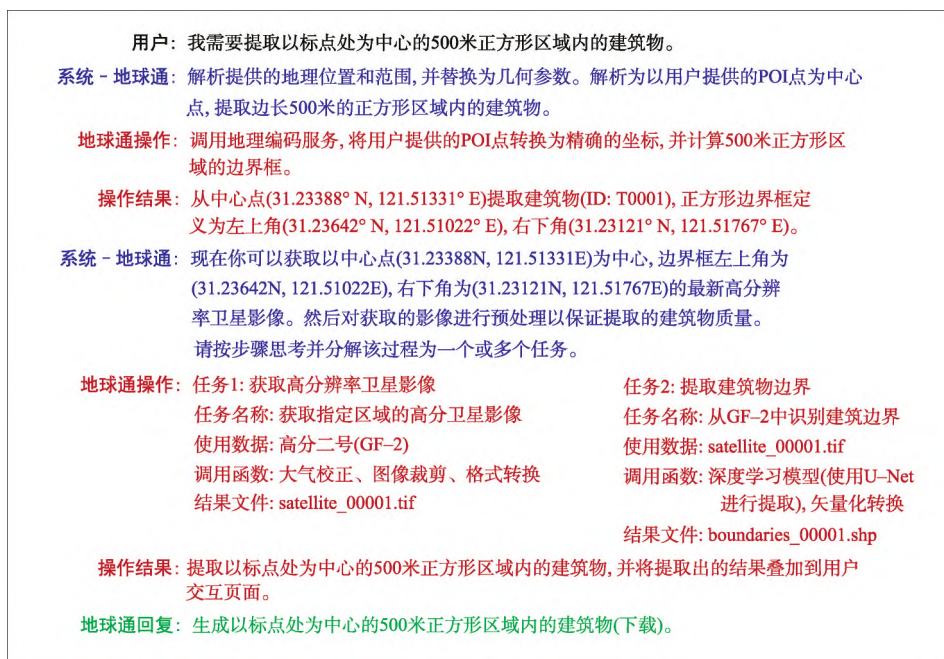


图7 地球通解题的多轮对话(部分)

Fig. 7 Multi-Round dialogue on EarthSage problem solving (partial)

将结果以叠加图层的形式显示在用户交互界面中,同时提供标准的地理空间数据格式(如.shp或.tif文件)供用户下载使用(图8)。从而基于平台功能间接构建智能体对物理世界的感知能力,为用户提供“观古今于须臾,抚四海于一瞬”空间认知服务;同时也能够调用平台提供的路径规划与导航接口,驱动用户在物理世界中行动,从而间接形成智能体的行动操控能力,系统实现界面如图8所示。

4.1 多模感知与行动操控模块实现

“地球通”多模感知模块通过调用全息地球平台承载的地球观测数据、地球模拟数据、经济统计数据 and 原住民知识资源API的方式实现,平台数据更新智能体能主动感知;同时多模感知模块还能调用平台提供的用户数据上传与语音对话功能,通过上传卫星遥感影像数据、包含空间位置内容的实时语音对话,从而模拟地理智能体视觉、听觉能力。行动操控模块主要基于智能中枢生成的任务列表,通过调用全息地球平台功能接口实现,如调用平台的数据更新接口对平台承载的数据内容进行修改,修改完成后智能体感知模块又能自动感知,从而形成“感知-决策-行动-感知”闭环能力;调用平台提供的路径规划与导航接口,帮助用户生成一条最优路径,指导用户在物理世界中的导航行进,同时全息地球平台中关于用户的实时位置数据不断改变,智能体多模感知模块又能够实时捕捉这种变化,如果发生偏航,能够反馈至智能中枢重新制定方案并生成任务列表。

4.2 智能中枢实现

“地球通”的实现关键在于智能中枢(大脑)的设计,本文采用大语言模型LLM组合GeoGPT构建智能中枢内核,连接多模感知模块、时序处理模块、

记忆存储模块、规划决策模块、行动操控模块,整体形成地理智能体大脑的架构,如图9所示。

GeoGPT为地球通提供了全方位的认知、推理和决策能力,通过处理多源异构数据,自动生成智能化的地理空间分析报告,预测环境变化,并在动态地理情境中实现实时决策。LLM负责地球通的自然语言处理和理解能力,而GeoGPT则专注于复杂地理现象的生成、解释和推理,确保地球通在多样化的地理空间环境中做出精确和合理的决策。这种多层次的认知与推理体系使地球通能够满足广泛的应用需求,从环境监测到城市规划,再到应急管理复杂场景。GeoGPT的实现以深度学习驱动的生成模型为核心,通过多模态数据压缩和分析技术,实现对遥感影像、无人机观测数据以及站点观测数据的实时处理,还能结合地球数据立方体“时空4维、属性 N 维”维度建模方法,通过多级网格剖分结合格元嵌入场技术,将不同传感器采集的矢量、栅格、格网等不同格式多源数据整合成64维嵌入向量,在实现高维数据压缩前提下,提取并保留各数据源之间的“关键共同信息”,实现多模态数据的高效压缩和互通。嵌入场技术和生成模型相结合,GeoGPT能够快速理解用户的自然语言输入进行空间内容的生成,同时还能够根据用户需求自动规划和执行工具链任务,从而调用全息地球平台功能以满足不同用户需求。例如,当用户提出“分析某地区的洪水风险”这一口语化问题时,GeoGPT能够自主完成从数据下载(气象、地形等)到特征提取,再到水文模型生成和洪水风险分析的完整流程。GeoGPT通过自主规划和执行,实现了智能化的数据处理过程,避免了用户在数据准备和工具配置上的繁琐操作,大幅提高了地理数据处理的自动



图8 地球通终端界面

Fig. 8 Interface of EarthSage client

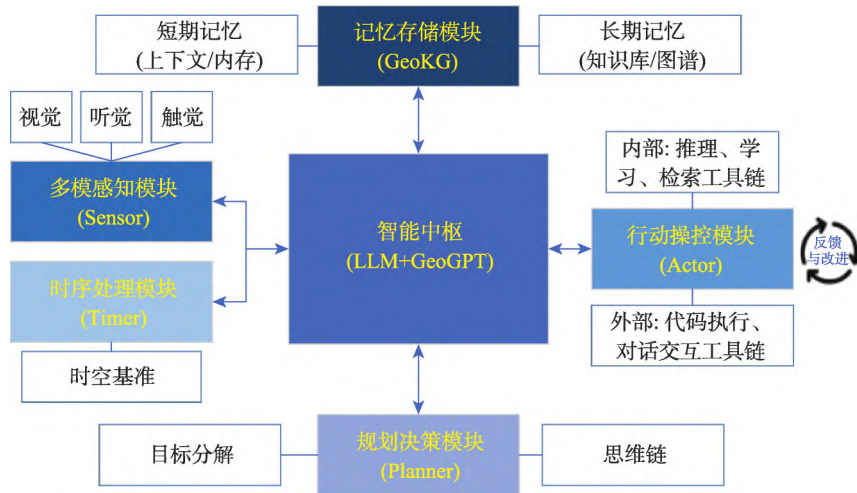


图9 地球通智能中枢架构
Fig. 9 Architecture of EarthSage brain

化水平和用户体验。

GeoKG作为智能中枢连接的记忆存储模块,承担着多源地理空间数据的整合与管理功能,支持系统进行决策制定和高级空间推理,增强 GeoGPT 地理现象解释与推理能力,从而避免大模型“幻觉”,为“地球通”提供精确、可信决策支持。GeoKG通过将来不同来源的异构数据,如卫星观测、无人机影像、地面监测数据和人类活动数据等,整合为统一的地理知识结构,构建一个涵盖自然系统(如土壤、空气、水体、生物等)和人类系统(如人口分布、社会经济信息、基础设施等)的综合性知识库,为传统 GIS 数据和模型分析提供语义增强层,实现时空数据以知识图谱结构的形式表示和存储,从而能够支持复杂的空间推理、扩展查询和多维数据分析。GeoKG的动态更新能力依赖多模态感知框架,通过实时整合来自卫星、无人机和地面站点的观测数据,确保知识体系保持持续更新。同时 GeoKG 与行动操控模块形成了双向反馈机制,进一步增强了系统的交互能力和智能化水平。

总之,智能中枢通过与各个模块的互动,特别是 GeoGPT 与 GeoKG 的双向互动机制,GeoKG 通过其强大的数据整合与语义推理能力,为地球通在处理复杂环境变化和支持决策分析时提供了稳固的知识基础,并通过与感知系统对接形成智能体实时响应能力; GeoGPT 从结构化的地理知识库中提取信息,并生成与特定地理场景相关的自然语言解释,确保了智能中枢能够根据不断变化的地理数据生成具有逻辑性和实用性的建议和报告,这种协同处理能力使得“地球通”不仅能支持用户的复杂地

理查询和空间推理需求,还能使人机交互更加直观和情境化,使得系统能够通过多层次的语义反馈大幅提升用户体验。

4.3 对话机器人系统集成

“地球通”采用对话机器人系统架构,通过自然语言处理与基于 GIS 工具链的查询响应系统,实现了用户与系统之间的高效交互(图 10)。该系统集成了智能对话引擎、实时服务系统和知识库,为用户提供精准的地理空间知识服务。通过这一架构,用户能够方便地以文本或语音输入与系统交互,系统则借助多层次模块实现从信息接收、解析到响应生成的全流程自动化,显著提升了信息查询与知识获取的效率与体验。

在系统的工作流程中,用户输入的文本或语音数据首先由自然语言处理引擎处理。自然语言处理模块将非结构化的用户输入转化为系统可处理的结构化数据。具体而言,意图分类器负责识别用户需求,而实体提取器进一步提取用户请求中的地理实体或具体位置。随后,对话管理代理运用策略学习和反馈机制,根据用户的需求生成适应性的行动计划。通过这种层次化处理,系统能够清晰解析用户的多样化输入,为后续的任务执行提供基础。

系统的行动计划由节点服务器执行。节点服务器根据预设行为模板和用户反馈,生成服务请求并调用相关 API 接口。实时服务系统则调用多种地理信息服务,包括瓦片地图服务、路线规划服务和地名检索服务等,为用户提供详细的地理数据和位置信息。通过用户的行为和偏好,系统能够优化

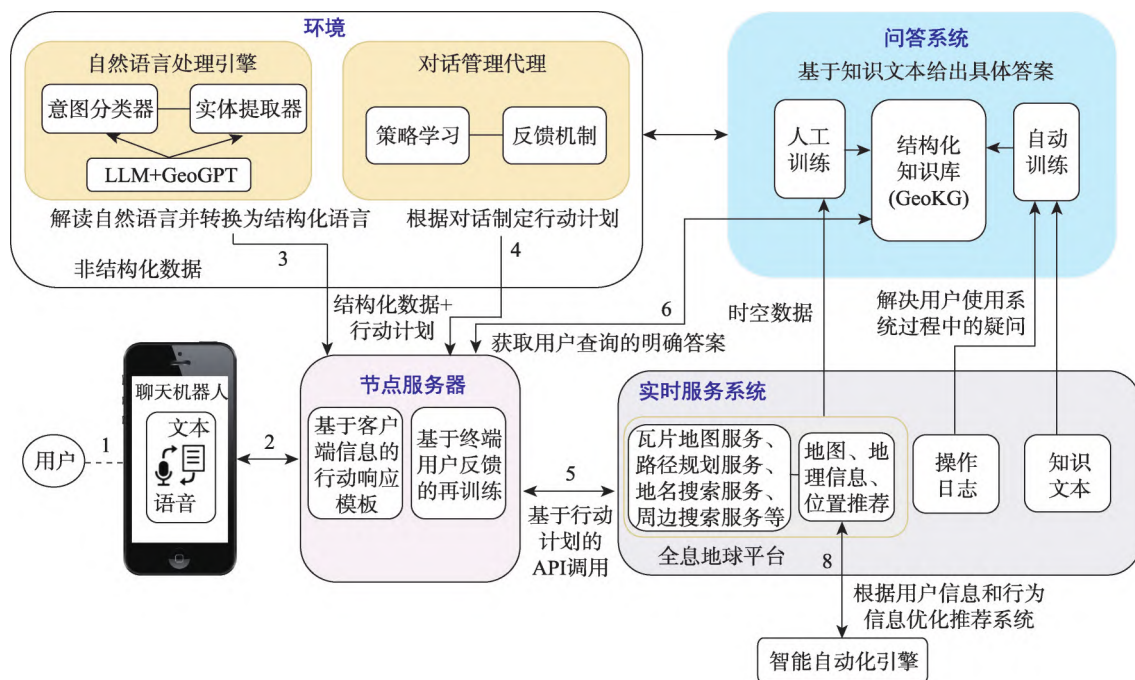


图 10 地球通聊天机器人系统架构

Fig. 10 Chatbot system architecture of EarthSage

信息推荐,提升互动体验和响应的个性化程度。这一实时服务系统确保了对话机器人在复杂的需求中能够提供及时准确的服务,自适应满足用户对不同场景下的地理信息服务与知识获取的需求。

此外,系统中集成的问答系统依托结构化知识库和自动化训练模块,不仅能解答常规查询,还可解决用户在使用过程中遇到的疑问。知识库通过人工训练和自动训练结合的方式不断丰富内容,使系统的对话能力逐步提升。系统通过这一知识库更新机制确保对话机器人能够响应用户的多样化需求并提供高效的地理空间知识服务。

总体而言,地球通的对话机器人系统架构通过多模块的紧密协作,实现了从用户输入到服务响应的自动化全流程。这一系统不仅适用于常规的地理数据查询,还能满足智慧城市管理、应急响应等复杂应用场景的需求,展现了对话机器人在地理知识服务中的广泛应用潜力。

5 结论

随着人工智能的快速发展,地理信息系统向地理智能体的演进已成为必然趋势。这一转变不仅是技术自然演进的结果,更是由复杂地理环境、动态变化以及多源异构数据处理需求的迫切性所驱动。相比传统的地理信息系统,地理智能体作为一

类结合最新AI技术的智能地理系统,具有更强的自主性和适应性,特别是在快速变化的环境中展现出更高的响应能力。自监督学习在这一演进过程中起到了关键作用,使地理智能体能够在无明确监督的情况下学习数据的内在结构,从而提升其在复杂数据处理中的泛化能力。这一特性帮助智能体适应动态地理环境,降低了对预定义模型的依赖,为智能体在不同场景中快速适应提供了技术基础。

在增强地理智能体适应性的同时,大型语言模型(LLM),尤其是像GeoGPT这样的生成式预训练模型,为系统的自然语言理解与生成能力带来了飞跃式提升。借助GeoGPT,地理智能体能够准确理解用户的自然语言输入,基于具体地理场景生成适当的分析和决策方案,使非专业用户也能够通过简单的语言指令完成复杂的地理数据分析与决策。例如,用户只需提出“分析某区域的洪水风险”,GeoGPT便可自动完成从数据采集、模型训练到分析评估的全过程,提供专业化的风险评估报告。GeoGPT的引入将推动地理智能体从被动响应向主动决策的转变,极大提升了用户体验和操作效率^[35]。

然而,地理智能体在应对复杂地理问题时仍面临诸多挑战,尤其是在与现实世界的交互和独立完成整套地学任务流的能力方面。本文提出的地理智能体的原型机“地球通”目前基于机器人问答系统,暂时缺乏直接与物理世界交互的能力。此外,

缺乏多层次规划能力也限制了“地球通”在长时间尺度或跨区域任务中的应用效果。尽管“地球通”能够满足短期任务需求,如道路规划或灾害响应,但在长期监测、多因素平衡任务(如城市扩展与资源管理)等复杂任务中,尤其是在依赖“地球通”直接与现实世界交互以完成地学任务时,现有的系统架构仍需进一步完善,以适应更复杂的决策需求。

尽管如此,通过对话机器人系统,“地球通”在现阶段已对地理智能体的构建思路进行了初步验证,“地球通”作为地理智能体的原型机,已具备展示本文所设计的地理智能体框架和主要功能模块的能力,为后续技术开发和应用奠定了重要的基础。未来,“地球通”将在以下方向取得突破:①进一步创新和优化交互模式,通过增强地理空间专业领域自然语言交互能力,实现更直观、高效的人机互动,降低专业门槛,让非专业用户通过与智能体对话就能轻松完成复杂的地理分析任务;②持续跟踪脑科学认知地图神经网络机理的研究,引入更高效的生成模型和知识图谱扩展技术,从而提升对多模态时空数据的处理能力,优化智能中枢推理与数智生成效率,实现从语言智能到空间智能的跨越;③构建模拟器与试验场,强化智能中枢与物理世界交互的物联感知与行动操控系统对接,形成“感知-认知-决策-行动”多模块有机组合的地理智能体系统。

地理信息系统向地理智能体的转变不仅体现了技术进步,也回应了现实世界对智能地理空间决策支持的迫切需求。随着具身智能、自监督学习、生成模型等先进人工智能技术的进一步融合^[36],地理智能体的能力将不断提升,不仅能够复杂动态环境中做出快速、准确的响应,还将推动智慧城市规划、生态环境监测和自然灾害应急响应等领域的智能化发展。这一演进标志着地理信息系统从静态的信息处理工具,向具备自主学习、实时适应和动态决策能力的方向转变,迈出了通往真正地理尺度空间智能的关键一步。

利益冲突: Conflicts of Interest

所有作者声明不存在利益冲突。

All authors disclose no relevant conflicts of interest.

作者贡献: Author Contributions

罗斌和刘文豪完成论文思路设计;韩嘉福,吴文周,李洪省,罗斌,刘文豪,吴进完成实验操作;刘文豪,罗斌,吴进参与论文的写作和修改。所有作者均阅读并同意最终稿件的提交。

LUO Bin and LIU Wenhao completed the design of the paper idea; HAN Jiafu, WU Wenzhou, LI Hongsheng, LUO Bin, LIU Wenhao, and WU Jin completed the experimental manipulation; LIU Wenhao, LUO Bin, and WU Jin participated in the writing and revision of the paper. All authors read and agreed to the submission of the final manuscript.

参考文献(References):

- [1] 陈述彭. 地理系统与地理信息系统[J]. 地理学报, 1991, 46(1): 1-7. [Chen S P. Geo-system and geo-information system[J]. Acta Geographica Sinica, 1991, 46(1): 1-7.] DOI: 10.11821/xb199101001
- [2] 陈健飞, 连莲. 地理信息系统导论[M]. 北京: 科学出版社, 2003. [Chen J F, Lian L. Introduction to geographic information systems[M]. Beijing: Science Press, 2003.]
- [3] 高松. 地理空间人工智能的近期研究总结与思考[J]. 武汉大学学报(信息科学版), 2020, 45(12): 1865-1874. [Gao S. A review of recent researches and reflections on geospatial artificial intelligence[J]. Geomatics and Information Science of Wuhan University, 2020, 45(12): 1865-1874.] DOI: 10.13203/j.whugis20200597
- [4] 张永生, 张振超, 童晓冲, 等. 地理空间智能研究进展和面临的若干挑战[J]. 测绘学报, 2021, 50(9): 1137-1146. [Zhang Y S, Zhang Z C, Tong X C, et al. Progress and challenges of geospatial artificial intelligence[J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(9): 1137-1146.] DOI: 10.11947/j.AGCS.2021.20200420
- [5] 苏奋振, 吴文周, 张宇, 等. 从地理信息系统到智能地理系统[J]. 地球信息科学学报, 2020, 22(1): 2-10. [Su F Z, Wu W Z, Zhang Y, et al. From geographic information system to intelligent geographic system[J]. Journal of Geo-Information Science, 2020, 22(1): 2-10.] DOI: 10.12082/dqxxkx.2020.190802
- [6] 李德仁, 朱欣焰, 龚健雅. 从数字地图到空间信息网格——空间信息多级网格理论思考[J]. 武汉大学学报(信息科学版), 2003, 28(6): 642-650. [Li D R, Zhu X Y, Gong J Y. From digital map to spatial information multi-grid: A thought of spatial information multi-grid theory[J]. Geomatics and Information Science of Wuhan University, 2003, 28(6): 642-650.] DOI: 10.3969/j.issn.1672-5603.2010.10.038
- [7] Liu Y, Chen W X, Bai Y J, et al. Aligning cyber space with physical world: A comprehensive survey on embodied AI [EB/OL]. 2024: 2407.06886. <https://arxiv.org/abs/2407.06886v7>.
- [8] Chen X L, He K M. Exploring simple Siamese representation learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2021: 15750-15758. DOI: 10.1109/cvpr46437.2021.01549.
- [9] Nagrani A, Yang S, Arnab A, et al. Attention bottlenecks for multimodal fusion[C]//Proceedings of the 35th International Conference on Neural Information Processing Systems. ACM, 2024: 14200-14213. DOI: 10.5555/3540261.3541349
- [10] 周成虎. 全空间地理信息系统展望[J]. 地理科学进展, 2015, 34(2): 129-131. [Zhou C H. Prospects on pan-spatial information system[J]. Progress in Geography, 2015, 34(2): 129-131.] DOI: 10.11820/dlkxjz.2015.02.001

- [11] 李德仁. 展望大数据时代的地球空间信息学[J]. 测绘学报, 2016, 45(4): 379-384. [Li D R. Towards geo-spatial information science in big data era[J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(4): 379-384.] DOI:10.11947/j.AGCS.2016.20160057
- [12] 黎夏, 叶嘉安, 刘小平, 等. 地理模拟系统: 元胞自动机与多智能体[M]. 北京: 科学出版社, 2007. [Li X, Ye J A, Liu X P, et al. *Geographic simulation system: Cellular automata and multi-agent*[M]. Beijing: Science Press, 2007.]
- [13] Ishikawa T. Spatial thinking, cognitive mapping, and spatial awareness[J]. *Cognitive Processing*, 2021, 22(Suppl 1): 89-96. DOI:10.1007/s10339-021-01046-1
- [14] 高俊, 曹雪峰. 空间认知推动地图学学科发展的新方向[J]. 测绘学报, 2021, 50(6): 711-725. [Gao J, Cao X F. The new development direction of cartography promoted by spatial cognition[J]. *Acta Geodaetica et Cartographica Sinica*, 2021, 50(6): 711-725.] DOI:10.11947/j.AGCS.2021.20210043
- [15] Rolls E T. The memory systems of the human brain and generative artificial intelligence[J]. *Heliyon*, 2024, 10(11): e31965. DOI:10.1016/j.heliyon.2024.e31965
- [16] Hoyt C T, Berrendorf M, Galkin M, et al. A unified framework for rank-based evaluation metrics for link prediction in knowledge graphs[EB/OL]. 2022: 2203.07544. <https://arxiv.org/abs/2203.07544v2>.
- [17] Ji S, Pan S, Cambria E, et al. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 20, 33(2): 494-514. DOI:10.1109/TNNLS.2021.3070843
- [18] Liu J J, Wang P, Shang Z Y, et al. IterDE: An iterative knowledge distillation framework for knowledge graph embeddings [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023, 37(4): 4488-4496. DOI:10.1609/aaai.v37i4.25570
- [19] Zhao T J, Wang S, Ouyang C J, et al. Artificial intelligence for geoscience: Progress, challenges, and perspectives[J]. *The Innovation*, 2024, 5(5): 100691. DOI:10.1016/j.xinn.2024.100691.
- [20] Jiao L C, Song X, You C, et al. AI meets physics: A comprehensive survey[J]. *Artificial Intelligence Review*, 2024, 57(9): 256. DOI:10.1007/s10462-024-10874-4
- [21] Epstein R A, Patai E Z, Julian J B, et al. The cognitive map in humans: Spatial navigation and beyond[J]. *Nature Neuroscience*, 2017, 20(11): 1504-1513. DOI:10.1038/nn.4656.
- [22] Fernandez-Leon J A, Acosta G G. Uncovering the secrets of the concept of place in cognitive maps aided by artificial intelligence[J]. *Cognitive Computation*, 2024, 16(5): 2334-2344. DOI:10.1007/s12559-022-10064-w
- [23] Gornet J, Thomson M. Automated construction of cognitive maps with visual predictive coding[J]. *Nature Machine Intelligence*, 2024, 6: 820-833. DOI:10.1038/s42256-024-00863-1
- [24] 何诗, 阴劼. 认知地图的地理学研究进展与展望[J]. 地理科学进展, 2022, 41(1): 73-85. [He S, Yin J. Progress and outlook of geographical research on cognitive maps [J]. *Progress in Geography*, 2022, 41(1): 73-85.]
- [25] Eyring V, Gentile P, Camps-Valls G, et al. AI-empowered next-generation multiscale climate modelling for mitigation and adaptation[J]. *Nature Geoscience*, 2024, 17: 963-971. DOI:10.1038/s41561-024-01527-w.
- [26] Jayaraman D, Grauman K. Learning to look around: Intellegently exploring unseen environments for unknown tasks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 1238-1247. DOI:10.1109/CVPR.2018.00135
- [27] G T, J F, and Sinapov J. Transferring Implicit Knowledge of Non-Visual Object Properties Across Heterogeneous Robot Morphologies[J]. *IEEE International Conference on Robotics and Automation*. DOI:10.1109/ICRA48891.2023.10160811
- [28] Gordon D, Kembhavi A, Rastegari M, et al. IQA: Visual question answering in interactive environments[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 4089-4098. DOI:10.1109/CVPR.2018.00430.
- [29] Shen W, Yang G, Yu A, et al. Distilled Feature Fields Enable Few-Shot Language-Guided Manipulation[C]//7th Annual Conference on Robot Learning. 20. <https://doi.org/10.48550/arXiv.2308.07931>.
- [30] Driess D, Xia F, Sajjadi M S M, et al. PaLM-E: an embodied multimodal language model[C]//Proceedings of the 40th International Conference on Machine Learning. Honolulu, Hawaii, USA: JMLR.org, 20. <https://doi.org/10.48550/arXiv.2303.03378>.
- [31] Huang W, Abbeel P, Pathak D, et al. Language Models as Zero-Shot Planners: Extracting Actionable Knowledge for Embodied Agents[J]. *International Conference on Machine Learning*. <https://doi.org/10.48550/arXiv.2201.07207>.
- [32] Torne M, Simeonov A, Li Z C, et al. Reconciling reality through simulation: A real-to-sim-to-real approach for robust manipulation[J]. *ArXiv e-Prints*, 2024: arXiv: 2403.03949. DOI:10.48550/arXiv.2403.03949
- [33] 罗斌, 任丽秋, 毛悦, 等. 全息地球数据立方体模型研究[J]. 地球信息科学学报, 2023, 25(7): 1282-1296. [Luo B, Ren L Q, Mao Y, et al. Research on the digital holographic earth data cube model[J]. *Journal of Geo-information Science*, 2023, 25(7): 1282-1296.] DOI:10.12082/dqxxkx.2023.230105
- [34] 黄勃, 吴申奥, 王文广, 等. 图模互补: 知识图谱与大模型融合综述[J]. 武汉大学学报(理学版), 2024, 70(4): 397-412. [Huang B, Wu S A, Wang W G, et al. KG-LLM-MCom: A survey on integration of knowledge graph and large language model[J]. *Journal of Wuhan University (Natural Science Edition)*, 2024, 70(4): 397-412.] DOI: 10.14188/j.1671-8836.2024.0040
- [35] Xiong H Y, Wang Z Y, Li X H, et al. Converging paradigms: The synergy of symbolic and connectionist AI in LLM-empowered autonomous agents[EB/OL]. 2024: 2407.08516. <https://arxiv.org/abs/2407.08516v5>.
- [36] Duan J, Yu S, Tan H L, et al. A Survey of Embodied AI: From Simulators to Research Tasks[J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2022, 6(2): 30-244. DOI:10.1109/TETCI.2022.3141105

■ 本文图文责任编辑: 蒋树芳 黄光玉