

引文格式: 庄晓东, 赵永峰, 胡振彪, 等. 大语言模型驱动的地理信息智能交互系统 [J]. 测绘科学, 2025, 50(3): 180-188. (ZHUANG Xiaodong, ZHAO Yongfeng, HU Zhenbiao, et al. Research on intelligent geographic information interaction method based on agent technology [J]. Science of Surveying and Mapping, 2025, 50(3): 180-188.) DOI: 10.16251/j.cnki.1009-2307.2025.03.020.

# 大语言模型驱动的地理信息智能交互系统

庄晓东<sup>1,2</sup>, 赵永峰<sup>1,2,3</sup>, 胡振彪<sup>1,2</sup>, 郭 杰<sup>4</sup>, 宗恒康<sup>1,2</sup>

- (1. 青岛市勘察测绘研究院, 山东 青岛 266033;
2. 青岛市海陆地理信息集成与应用重点实验室, 山东 青岛 266033;
3. 青岛市西海岸基础地理信息中心有限公司, 山东 青岛 266555;
4. 山东省烟台生态环境监测中心, 山东 烟台 264000)

**摘 要:** 针对传统的 GIS 系统通常需要通过复杂的图形界面和固定的任务流程来响应用户的操作的问题, 对大语言模型驱动的地理信息智能交互系统进行了研究, 即充分利用大语言模型的能力, 使得 GIS 系统能够真正“理解”人类的意图, 并自主规划完成任务。基于 Agent 技术实现 GIS 智能交互, 利用 JSON Schema 规范实现了模糊自然语言到 GIS 专业工具的映射, 基于 ReAct 模式实现了 GIS 工具多轮迭代组合调用, 采用名称坐标分离策略实现了空间信息感知, 使用 RAG 技术实现了本地知识库查询与互联网信息检索。本文打造出一款可应用于实际场景中的青岛市智能出行小助手, 并进行实验和大语言模型对比评估。

**关键词:** 人工智能; 智能体; 大语言模型; 智能地理信息系统

**【中图分类号】** P208

**【文献标志码】** A

**【文章编号】** 1009-2307(2025)03-0180-09

## Research on intelligent geographic information interaction method based on agent technology

ZHUANG Xiaodong<sup>1,2</sup>, ZHAO Yongfeng<sup>1,2,3</sup>, HU Zhenbiao<sup>1,2</sup>, GUO Jie<sup>4</sup>, ZONG Hengkang<sup>1,2</sup>

- (1. Qingdao Surveying & Mapping Institute, Qingdao, Shandong 266033, China;
2. Key Laboratory of Sea-land Geographics Information Integration and Application of Qingdao, Qingdao, Shandong 266033, China;
3. West Coast Geomatics Center of Qingdao Co., Ltd., Qingdao, Shandong 266555, China;
4. Shandong Provincial Yantai Ecological Environment Monitoring Center, Yantai, Shandong 264000, China)

**Abstract:** Traditional GIS systems have certain limitations in their interaction methods, typically featuring complex graphical interfaces and fixed task workflows to respond to user operations. To address this issue, this paper investigates a Geographic Information Intelligent Interaction System driven by a large language model, leveraging the capabilities of the large language model to enable GIS systems to truly “understand” human intentions and autonomously plan and complete tasks. This paper implements GIS intelligent interaction based on Agent technology, realizes the mapping from fuzzy natural language to GIS professional tools based on the JSON Schema specification, implements multi-round iterative combination calls of GIS tools based on the ReAct pattern, adopts a name-coordinate separation strategy for spatial information perception, and uses RAG technology for local knowledge base queries and internet information retrieval. This paper successfully develops a practical application, the Qingdao Smart Travel Assistant, and conducts experiments and comparative evaluations with different large language models.

**作者简介:** 庄晓东(1986—), 男, 山东烟台人, 高级工程师, 硕士, 主要研究方向为地理信息系统研发与空间数据分析。

**E-mail:** zhuangxiaodong@foxmail.com

**收稿日期:** 2024-08-01

**基金项目:** 泰山产业领军人才工程专项(新型时空信息基础设施构建关键技术研发与应用)

**通信作者:** 赵永峰 高级工程师 E-mail: 254837341@qq.com

This innovative approach brings new possibilities to intelligent GIS interactions, providing users with a more intelligent and flexible GIS system operation experience.

**Keywords:** artificial intelligence; agent; large language model; intelligent geographic information system

## 0 引言

通常的地理信息系统(geographic information system, GIS)以电子地图或影像为底图,并在界面上提供功能列表。用户指令主要通过鼠标和键盘完成,包括鼠标点击、拖拽和滚轮等方式实现地图的交互,通过键盘录入信息等。用户按照系统既定的规则和流程执行操作,以获取预期结果。这种传统 GIS 系统设计依赖用户熟练掌握特定的界面和指令语言,对用户的专业知识和操作经验提出了一定要求,导致学习曲线较陡峭,限制了系统的普适性和用户友好性。

人与人之间的信息交互通常通过自然语言完成。在 GIS 系统设计方面,如何基于自然语言进行人机交互一直是许多学者研究和探索的目标。然而,在大语言模型出现之前,类似的研究方法通常依赖于基于规则的穷举和匹配<sup>[1-2]</sup>,这在面对更为复杂的现实环境时存在一定的局限性,比如无法处理不规范的表达或者需要多轮工具调用的场景。

Vaswani 等人于 2017 年提出了基于自注意力机制的 Transformer 深度学习架构。该架构具有强大的并行处理能力和能够捕捉长距离依赖关系的优势,因而成为现代大语言模型的基石<sup>[3]</sup>。随后,OpenAI 团队在 2018 年发布了第一代 GPT 模型,当时的大语言模型仍需要针对下游任务进行微调<sup>[4]</sup>。

OpenAI 于 2021 年发布了 GPT-3,引入了 Few-shot 的全新应用模式<sup>[5]</sup>。在使用诸如 ChatGPT 等大语言模型时,用户只需构建适当的提示词(Prompt),即可执行翻译、提取、分类、创作等任务,这种应用模式为 GIS 系统的智能化带来了新的可能性。很多学者评估了 GPT-4 等大语言模型对地理基础知识的掌握情况及其将这些知识用于解释推理(如地理空间分析)的能力<sup>[6-8]</sup>。然而,原生大语言模型的回答存在不受控制的问题,难以直接应用于需要精确指令的业务场景。

2023 年,以 AutoGPT<sup>[9]</sup>、AutoGen<sup>[10]</sup>为代表的智能体(Agent)技术成为大语言模型落地场景的重要突破口。Agent 是一种以大语言模型为核心,具备环境感知、决策制定和行动执行能力的高级

智能实体。与传统 AI 相比,Agent 能够针对特定目标进行自主规划和策略实施,利用外界反馈和自我思考实现任务自动化<sup>[11]</sup>。将 Agent 技术应用于业务场景中,已成为当前研究的热点。在 GIS 领域,一些学者将部分的 Agent 能力应用于地图制图<sup>[12]</sup>、空间分析<sup>[13-14]</sup>等领域,取得了很好的应用成效,但其主要着眼于 GIS 工具调用,在 GIS 系统层级方面的研究较少。

本文设计和实现了一种基于自然语言的 GIS 智能交互系统,大语言模型充当用户需求与 GIS 系统之间的中介,将用户需求转化为 GIS 系统能够理解的指令。在此过程中,大语言模型的智能特性使得用户可以使用更加灵活、自然的语言来操作 GIS 系统。本文将 Agent 技术融入到 GIS 系统的架构设计、功能开发以及业务流程,可以为 GIS 系统提供任务规划、推理、记忆、互联网检索和本地知识库查询等扩展功能。

## 1 GIS 智能交互系统架构

本文设计和实现的地理信息智能交互系统整体架构如图 1 所示,由下而上分为数据层、模型层、应用层和交互层的四层架构。

数据层由多源数据构成。除了传统的 GIS 矢量/栅格数据外,本地知识库和互联网检索的信息可以作为知识补充,增强 GIS 系统的业务能力。

模型层由大语言模型驱动,基于 Agent 技术实现感知、规划、行动、反思和记忆的能力,可以接受用户需求,结合当前状态,将一个总的目标任务拆分为一系列步骤分步执行,在每一步执行完成后反思执行结果。模型的记忆能力可以实现对应用场景整个生命周期的掌控。

应用层包含后端 GIS 能力,包括但不限于地名地址查询、空间查询、空间分析和统计、SQL 查询地理数据库等工具。这些工具基于大语言模型的编排执行,并由前端地图引擎实现对 GIS 执行结果的展示,其灵活性有助于 GIS 系统由单一场景到多场景应用的拓展。

交互层相较于传统 GIS 系统,其优势体现为对话式的 GIS 系统操作,因此系统 UI 界面不需要复杂的菜单和工具,仅凭借对话窗口实现任务执行和知识检索智能问答。

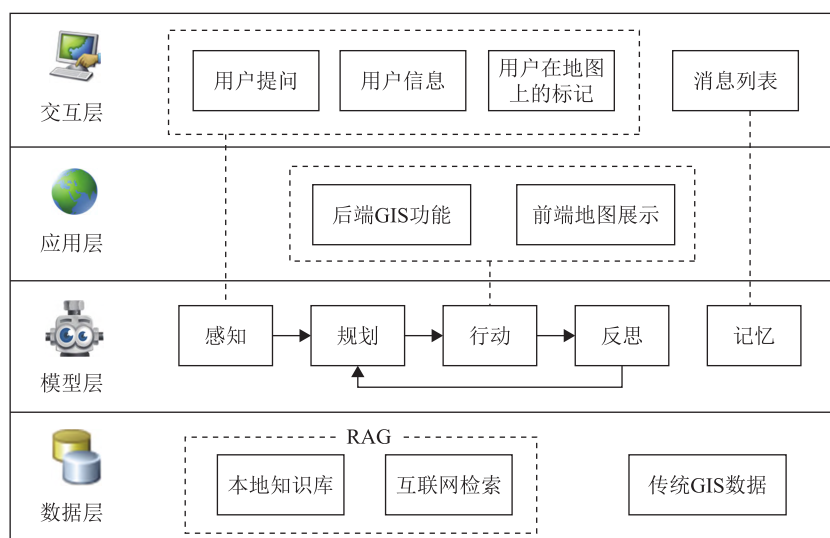


图 1 GIS 智能交互系统架构

Fig. 1 The Architecture of GIS Intelligent Interaction System

## 2 基于 Agent 技术实现 GIS 智能交互

本文应用 Agent 的 Tool Calling 能力与 GIS 系

统后端工具接口进行对接, 流程如图 2 所示。

在与 Agent 进行交互之前, 必须提供符合 JSON Schema 规范的工具描述文档, 确保 Agent

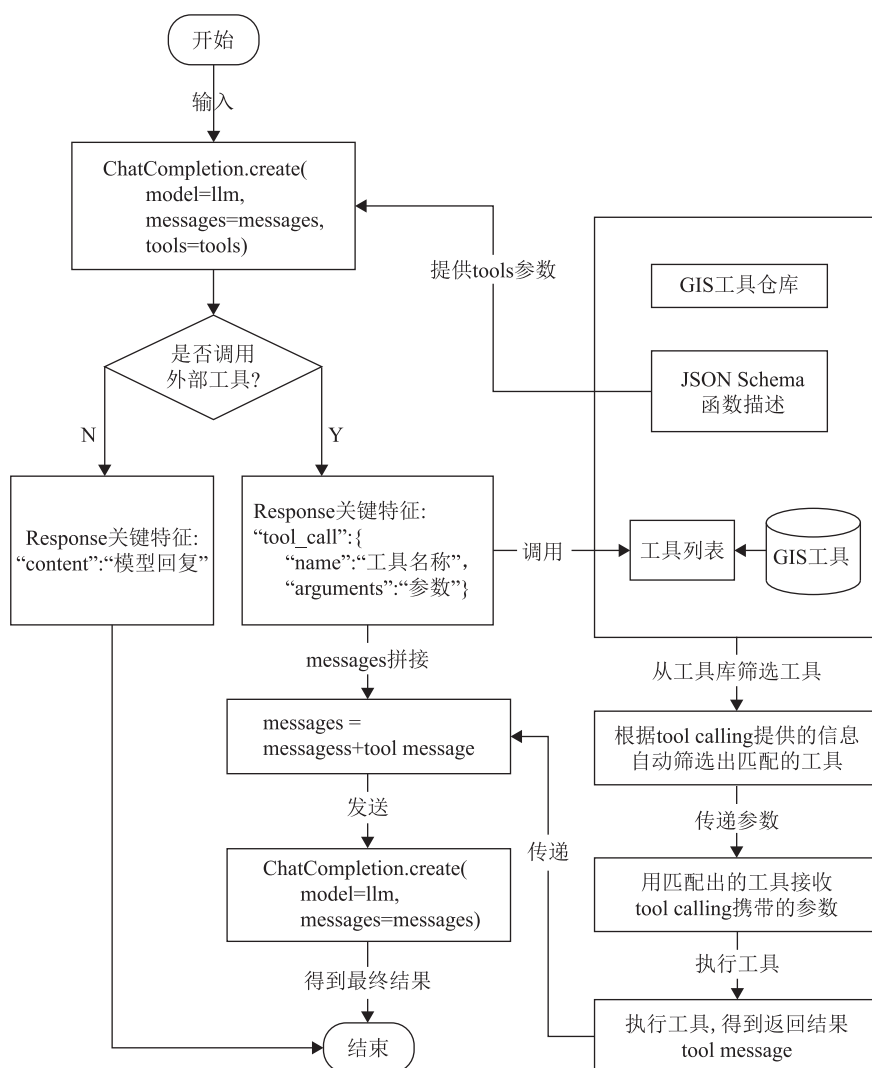


图 2 工具调用流程图

Fig. 2 Tool Calling Flowchart

对可用工具有深入且全面的理解。JSON Schema作为一种元数据标准,用于描述JSON数据的格式与结构,它支持对对象属性类型、数组长度、数值范围以及字符串模式等进行灵活验证。利用JSON Schema工具描述,Agent能够将用户的非规范自然语言需求转化为对应的工具函数调用。以路径规划为例,JSON Schema工具描述如图3所示。用户可能使用如“找一条青岛市政府去万象城的步行路线”或者“我想从青岛市政府溜达到万象城”等自然语言指令。大语言模型能够理解这些指令,并将它们映射为包含“tool\_call”字段的信息结构,其中包含函数名称(“name”:“path”)和参数列表(“start”:“青岛市政府”,“end”:“万象城”,“mode”:“步行”)。

```
{
  "name": "path",
  "description": "规划路线,帮助用户想去的地方",
  "parameters": {
    "type": "object",
    "properties": {
      "start": {
        "type": "string",
        "description": "出发地,可以接受用户在地图上标记的点位",
      },
      "end": {
        "type": "string",
        "description": "目的地,可以接受用户在地图上标记的点位",
      },
      "mode": {
        "type": "string",
        "description": "出行方式",
        "enum": ["驾车", "骑行", "步行", "公共交通"],
      },
    },
    "required": ["start", "end"],
  },
}
```

图3 路径规划工具描述JSON Schema示例

Fig.3 Example of JSON Schema for Path Planning  
Tool Description

以上流程将多样化、不规范的自然语言操作指令解析为GIS系统可执行的操作命令。这不仅实现了与GIS系统后端工具接口的无缝对接,而且从用户的角度来看,通过自然语言与GIS系统进行交互变得切实可行。

在处理用户需求的过程中,单一工具往往难以提供全面的解决方案。因此,本文引入了ReAct(reasoning and acting)模式,通过系统化地调用多个工具,实现任务的高效完成<sup>[15]</sup>。ReAct整体运作流程如图4所示,其核心理念在于赋予Agent观察环境的能力,以收集必要的信息,通过循环执行“思考(推理)-行动(执行动作)-观察(反馈结果)”的流程,Agent能够在任务执行过程中进行动态的推理与决策,从而显著提升任务解决的精确度。

ReAct模式的提示词模板如图5所示,它通过递归发送指令的方式,构建了连贯的上下文环境。这种环境使得Agent能够根据任务的完成进度进行实时规划,选择并分步执行所需的工具,以实现任务的全面完成。例如,当用户提出需求:“我想开车从青岛市勘察测绘研究院开车去新业广场,把车放下,坐公交去万象城,找一家咖啡店喝咖啡”,系统将依次进行两次路径规划和一次公共交通查询,并将这三轮工具调用的结果在地图上进行展示。

Agent本身不具备对空间信息的理解能力,如何能够“感知”用户在地图上的标记,也是需要解决的关键问题之一。本文采用标记名称和坐标分离的策略。标记名称和坐标通过键值对关联,在向Agent发送指令时,仅传递标记名称,并声明用户标记可以作为工具参数使用。这一传递过程可以通过两种途径实现,分别是通过系统提示词和记忆列表。

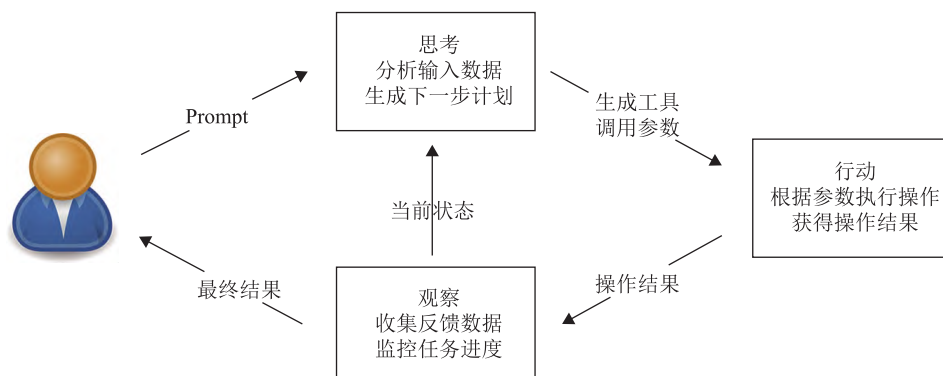


图4 ReAct流程图

Fig.4 ReAct Flowchart



作为专业的智能化GIS系统，你需要解决用户问题。  
你可以使用以下工具：

工具1: <工具1的描述>  
工具2: <工具2的描述>  
.....

**问题:** 你需要回答的问题  
**思考:** 你应该总是思考要做什么  
**行动:** 要采取的行动，应该是以下之一 [ 工具1, 工具2, ..... ]  
**行动输入:** 行动的名称  
**观察:** 行动的结果  
..... (这个**思考/行动/行动输入/观察**可以重复N次)  
**思考:** 我现在知道最终答案了  
**最终答案:** 对原始输入问题的最终答案

开始!

问题: <用户输入的问题>

图 5 ReAct 提示词模板

Fig. 5 ReAct Prompt Template

如果使用在记忆列表传入标记名称，为了避免添加标记的消息被“遗忘”，需要将此消息放置于每一轮任务的前一个位置。然而，这样做可能导致对话的不连贯，容易引起“困惑”。为了解决这一问题，本文的选择是将标记信息放入系统提示词中。这种策略有助于保持对话的流畅性，同时确保系统配置信息与用户交互的历史不冲突，提高了对话的可解释性和一致性。当系统接收到大语言模型返回的信息，并识别到标记名称作为参数时，系统通过键值对映射到对应的标记坐标，并将坐标作为真正的参数传入工具进行调用。通过这种方式，系统能够以灵活的自然语言方式控制精准的坐标信息，实现大语言模型对于地图标记的“感知”。

### 3 基于 RAG 技术的智能问答能力构建

检索增强生成(retrieval augmented generation, RAG)是一种使用外部知识增强 AI 回答准确度的技术。RAG 的一个显著优势在于它可以减少由于模型自身知识限制而产生的幻觉<sup>[16]</sup>。

基于 RAG 技术的本地知识库查询流程如图 6 所示。通常，GIS 系统都会有特定业务场景的知识文档。通过将这些文档进行词嵌入(Embedding)处理并存储于向量数据库中，形成本地知识库。词嵌入是将高维的文本信息映射到低维度的实数连续向量空间，基于分布式假设捕捉语义信息的技

术，因此本地知识库中的索引即为文本分块的向量。

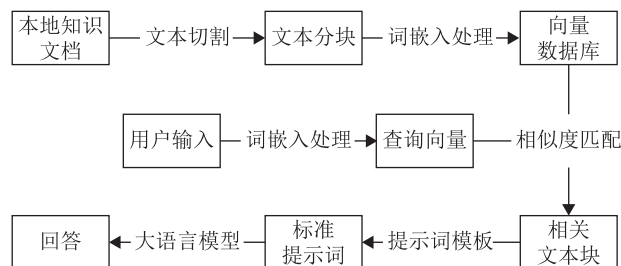


图 6 基于 RAG 技术的本地知识库查询

Fig. 6 Local Knowledge Base Query Based on RAG Technology

当用户输入消息后，系统可以利用向量匹配技术实现语义相似度匹配，从本地知识库中检索出与问题相关的文本信息；将这些检索到的信息与用户的原始问题相结合，形成丰富的提示信息。这样，大型语言模型便能够在包含外部信息的提示中进行上下文学习(in-context learning)，并据此生成准确、可靠的答案。

互联网检索的 RAG 流程更为简单，它无需进行复杂的向量转换和语义相似度匹配，只需利用搜索引擎的 API 在线检索相关文本即可。整体流程如图 7 所示。

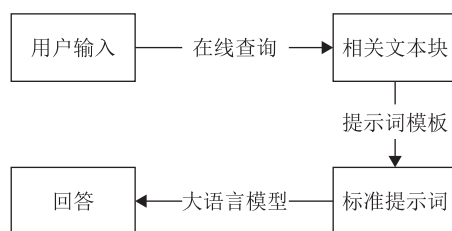


图 7 基于 RAG 技术的互联网信息检索

Fig. 7 Internet Information Retrieval Based on RAG Technology

## 4 案例研究

作为本文技术方法的实现案例，我们开发了“青岛市智能出行小助手”系统，整体架构如图 8 所示。此系统采用前后端分离的架构，前端使用 Mapbox 地图引擎进行地图方面的交互和展示，后端采用 Django 框架封装 GIS 工具，大语言模型选用清华智谱 AI 的 GLM-4 在线接口。在外部工具方面，我们针对出行辅助的需求，提供了地名地址查询、周边查询、路线规划(含公交、驾车、步行、骑行)、地图测量等 GIS 功能，后台调用了高德地图 API 实现；我们将青岛市景点信息整理成为本地知识库，采用 FAISS 向量数据库存储，基于 RAG 技术实现景点信息检索问答；时事信息检索和天气信息检索基于 Serp API 实现。

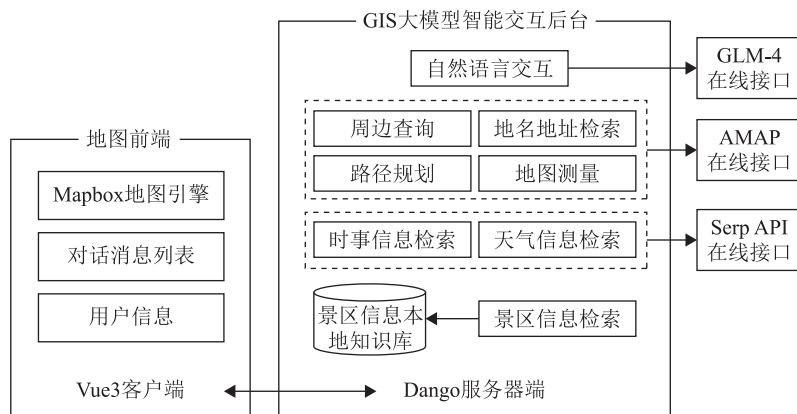


图8 “青岛市智能出行小助手”系统架构

Fig. 8 System Architecture of the “Qingdao Smart Travel Assistant”

系统界面如图9所示。左侧为自然语言交互窗口，右侧为地图区域。任务执行结果可以通过文字和地图同时进行反馈。

表1~表3是几个出行辅助任务的经典案例，涉及到多轮GIS工具调用、地图标记感知、在线检索、本地知识库检索等能力。

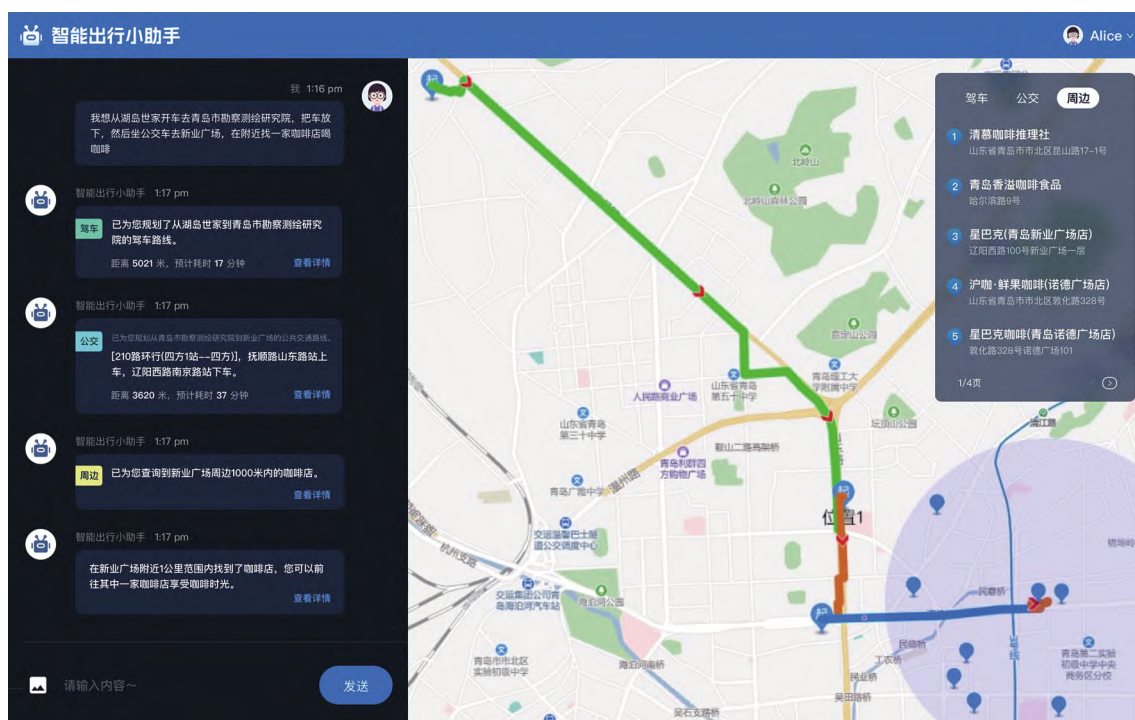


图9 “青岛市智能出行小助手”系统界面

Fig. 9 System Interface of the “Qingdao Smart Travel Assistant”

表1 在线检索示例

Tab. 1 Example of Internet Information Retrieval

用户	明天适合户外活动吗?
思考: 我需要查询青岛市明天的天气。	
行动: web_search	
行动输入: “明天青岛的天气”	
观察: 29日(明天). 晴转多云. 25℃/19℃. 3-4级	
思考: 我现在知道了最终答案	
最终答案: 明天适合户外活动。虽然天气会从晴转多云, 但最高温度25℃, 适宜户外活动。最低温度为19℃, 也不会太冷。	

表 2 本地知识库查询示例

Tab. 2 Example of Local Knowledge Base Query

用户	水准原点在哪个景区内? 由什么单位负责维护?
	思考: 这是一个关于景点信息问题, 需要借助工具 <code>get_qd_scenic_info</code> 来回答。
	行动: <code>get_qd_scenic_info</code>
青岛市智能出行小助手	行动输入: "水准原点在哪个景区内? 由什么单位负责维护?"
	观察: 观象山位于中国山东省青岛市市南区, 是观赏青岛市容及海景的佳处, 观象山顶处还有中国人民解放军总参测绘局于 50 年代所建的全国水准原点, 由青岛市勘察测绘研究院维护。
	思考: 我现在知道了最终答案
	最终答案: 水准原点位于观象山景区内, 由青岛市勘察测绘研究院负责维护。

表 3 多轮 GIS 工具调用(含地图标记感知)示例

Tab. 3 Example of Multi-Turn GIS Tool Calling(Including Map Annotation Perception)

用户	我想从这里开车去观象山, 在附近找一家加油站。
	思考: 我需要查找位置 1 到观象山的驾车路线, 然后搜索观象山周边的加油站。
	行动: <code>path</code>
	行动输入: { <code>start</code> : "位置 1", <code>end</code> : "观象山", <code>mode</code> : "驾车" }
	观察: { <code>response</code> : "路径规划已完成。", <code>path_info</code> : <驾车路径的详细信息 JSON, 包含路线坐标串、距离和耗时, 可在地图上展示> }
青岛市智能出行小助手	思考: 我需要查询观象山周边的加油站, 使用 <code>find_nearby_places</code> 工具
	行动: <code>find_nearby_places</code>
	行动输入: { <code>center</code> : "观象山", <code>place</code> : "加油站", <code>radius</code> : 1000 }
	观察: { <code>response</code> : "已为您查询到观象山周边 1000 米内的加油站。", <code>path_info</code> : <查询结果的详细信息 JSON 列表, 包含加油站名称和坐标, 可在地图上展示> }
	思考: 我现在知道了最终答案
	最终答案: 任务已完成。

由表 1~表 3 示例可以看出, 在系统知识和能力范围内的智能出行辅助方面的问题, 系统能够做出较好的反馈。然而, 用户在不知悉系统预设功能的前提下, 也会提出诸如“帮我查找附近优惠力度最大的加油站”这样的问题, 系统并不具备查询加油站优惠信息的功能, 理应给出“我无法获取加油站的优惠信息”这样的答复, 而大语言模型的“幻觉”问题可能会导致系统在周边加油站中任选一个回复给用户。此类问题是开放式问答系统的弊端, 需要靠提示词工程和大语言模型能力两方面的提升来缓解。

作为自然语言交互的基础底座, 大语言模型自身的精准度直接影响到系统的可靠性。本文对四种先进的大型语言模型——GPT-4、GPT-3.5、GLM-4 以及 ChatGLM3-6b——进行了比较分析, 评估它们作为 Agent 内核时, 规划调用 GIS 工具时的准确性。模型的选择综合考虑了国内外、开源和商业模型的多样性, 以确保全面评估不同类型模型在 GIS 任务中的性能表现。在测试数据生

成方面, 我们对用户实际输入的指令进行了筛选, 并利用 ChatGPT 进行数据增强以扩充数据量, 最终构建了包含 1 077 条测试数据的数据集, 涵盖了地址搜索、周边查询、路径规划、距离测量等 GIS 功能。测试数据中包含了多轮工具调用和含地图标记参数的复杂场景。取两条测试数据示例如表 4 所示。

可以看出, 每一轮函数调用的目标值都有函数名称和参数。大语言模型当前轮次返回的工具调用的 JSON 结果中的函数名称和参数与目标值相同, 则视为调用成功。测试结果如表 5 所示。可以看出, 随着任务复杂度的增加(工具调用轮次的增多、地图标记的加入), 工具调用的准确度会有所下降。在 3 种商业模型中, 即 GPT-4、GPT-3.5 和 GLM-4, 均展现出了较高的准确度。然而, 开源的 ChatGLM3-6b 模型由于参数量的限制, 导致其在处理多轮指令时的性能不佳, 因此在本文的任务中不具备可用性。



表4 测试数据示例

Tab. 4 Examples of Test Data

类型	用户消息	首轮目标值	二轮目标值	三轮目标值
1	单轮工具调用	<pre>{'role': 'assistant', 'tool_call': {'name': 'map_locate', 'arguments': {'location': '五四广场'}}}</pre>		
2	三轮工具调用(含地图标记参数)	<pre>{'role': 'assistant', 'tool_call': {'name': 'path', 'arguments': {'start': '位置1', 'end': '中山公园', 'mode': '驾车'}}}</pre>	<pre>{'role': 'assistant', 'tool_call': {'name': 'path', 'arguments': {'start': '中山公园', 'end': '植物园', 'mode': '驾车'}}}</pre>	<pre>{'role': 'assistant', 'tool_call': {'name': 'find_nearby_places', 'arguments': {'distance': 0.5, 'location': '植物园', 'place': '停车场'}}}</pre>

表5 模型 GIS 工具调用对比

Tab. 5 Model Comparison for GIS Tool Calling

指令类型		GPT-4	GPT-3.5	GLM-4	ChatGLM3-6b
单轮工具调用	成功率	100.0	99.3	99.0	88.1
	首轮成功率	98.2	98.9	97.8	87.0
两轮工具调用	两轮成功率	97.5	95.3	96.7	12.7
	首轮成功率	97.3	97.3	98.2	84.5
三轮工具调用	两轮成功率	96.4	92.7	95.5	11.0
	三轮成功率	93.6	88.2	91.0	7.3
	首轮成功率	95.2	88.6	96.2	89.5
三轮工具调用(含地图标记参数)	两轮成功率	92.4	87.6	91.4	16.2
	三轮成功率	92.4	79.0	87.6	12.4
	首轮成功率				

## 5 结束语

本文研究了大语言模型驱动的 GIS 智能交互系统的关键技术方法。实验结果表明,这种新型的交互方法能够简化 GIS 系统的操作界面,降低用户的使用门槛,并具备一定的容错性,使用户指令更加灵活;借助 Agent 技术, GIS 系统能够自主规划完成复杂任务,是智能 GIS 系统的一种可行的实现方法。

未来,我们将进一步优化这一交互方式,并探索大语言模型在 GIS 系统中更多的应用领域。继续研究的方向如下:将 Agent 接入更复杂、更专业的 GIS 业务系统中,例如城市规划、土地管理等领域,并通过分层级的模块组装进行业务流程的自动规划;在感知层面,借助多模态技术,将当前地图信息作为输入的一部分,使得 Agent 能够更全面地感知当前环境状态;目前的研究基于消息列表实现了短期记忆,而长期记忆面

临知识总结、构建索引等难题,未来将尝试借助知识图谱的 RAG 等技术进行有针对性的研究突破。

## 参考文献

- [1] 刘晓.面向高德地图的自然语言接口语义解析系统设计与实现[D].南京:南京师范大学,2015.(LIU Xiao. Design and implementation of natural language interface semantic parsing system for gaode map[D]. Nanjing: Nanjing Normal University, 2015.)
- [2] 徐娟.面向地图搜索领域的自然语言句子查询语义意图理解方法研究[D].上海:复旦大学,2014.(XU Juan. Research on semantic intention understanding method of natural language sentence query for map search field[D]. Shanghai: Fudan University, 2014.)
- [3] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30: 5998-6008.
- [4] RADFORD A, NARASIMHAN K. Improving language understanding by generative pre-training[Z]. [S. l.: s. n.],



- 2018.
- [5] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners[J]. *Advances in neural information processing systems*, 2020, 33:1877-1901.
- [6] Mai G, Huang W, Sun J, et al. On the opportunities and challenges of foundation models for geospatial artificial intelligence[J]. *arXiv preprint arXiv*: 2304.06798, 2023.
- [6] MAI Gengchen, HUANG Weiming, SUN Jin, et al. On the opportunities and challenges of foundation models for geospatial artificial intelligence[EB/OL]. [2024-12-01]. <https://arxiv.org/abs/2304.06798v1>.
- [7] ROBERTS J, LÜDDECKE T, DAS S, et al. GPT4GEO: How a language model sees the world's geography [EB/OL]. 2023; 2306. 00020. <https://arxiv.org/abs/2306.00020v1>.
- [8] MOONEY P, CUI W, GUAN B, et al. Towards Understanding the Spatial Literacy of ChatGPT[C]// *ACM SIGSPATIAL International Conference, [S.l. : s. n.]*, 2023.
- [9] YANG Hui, YUE Sifu, HE Yunzhong. Auto-GPT for online decision making: Benchmarks and additional opinions[EB/OL]. 2023; 2306. 02224. <https://arxiv.org/abs/2306.02224v1>.
- [10] WU Qingyun, BANSAL G, ZHANG Jieyu, et al. AutoGen: Enabling next-gen LLM applications *via* multi-agent conversation[EB/OL]. 2023; 2308. 08155. <https://arxiv.org/abs/2308.08155v2>.
- [11] WANG Lei, MA Chen, FENG Xueyang, et al. A survey on large language model based autonomous agents[J]. *Frontiers of Computer Science*, 2024, 18(6):186345.
- [12] ZHANG Yifan, HE Zhengting, LI Jingxuan, et al. MapGPT: an autonomous framework for mapping by integrating large language model and cartographic tools [J]. *Cartography and Geographic Information Science*, 2024, 51(6):717-743.
- [13] ZHANG Yifan, WEI Cheng, WU Shangyou, et al. GeoGPT: Understanding and processing geospatial tasks through an autonomous GPT[EB/OL]. [2024-09-09]. <https://arxiv.org/abs/2307.07930v1>.
- [14] LI Z, NING H. Autonomous GIS: the next-generation AI-powered GIS[J]. *International Journal of Digital Earth*, 2023, 16(2):4668-4686.
- [14] LI Zhenlong, NING Huan. Autonomous GIS: the next-generation AI-powered GIS[J]. *International Journal of Digital Earth*, 2023, 16(2):4668-4686.
- [15] YAO Shunyu, ZHAO Jeffrey, YU Dian, et al. React: Synergizing reasoning and acting in language models [EB/OL]. 2022; 2210. 03629. <https://arxiv.org/abs/2210.03629v3>.
- [16] LEWIS P, PEREZ E, PIKTUS A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks[J]. *Advances in Neural Information Processing Systems*, 2020, 33:9459-9474.