Self-evolving Spatial Intelligence - Integrating Large Language Models with GIS for Automated Geospatial Knowledge Discovery and Decision Support

Haotian Cheng*
Shandong Normal University, Jinan 250000, China
*Corresponding email: 15206705105@163.com

Abstract

Recent breakthroughs in natural language understanding, reasoning, and code generation have opened a new path for the intelligent transformation of geographic information systems (GIS). Existing large language models (LLM) -driven GIS methods. However, largely remain task-specific and static, lacking the ability to accumulate knowledge and evolve autonomously. This paper proposes a self-evolving spatial intelligence framework that deeply integrates LLM with GIS for automated geospatial knowledge discovery and decision support. The framework begins with the construction of a geospatial knowledge graph, extracting, organizing, and continuously updating spatial knowledge from multi-source heterogeneous data through a reflexion-enhanced self-learning loop. Building on this, an autonomous spatial analysis agent converts natural-language intents into executable geospatial workflows through tool calling and multi-agent collaboration. Finally, a knowledge-augmented decision-support layer produces interpretable, evidence-based recommendations for spatial planning and decision-making. Together, these three components form a closed loop of knowledge discovery, decision support, and self-evolution, moving GIS from a passive analysis tool toward a continuously self-improving spatial intelligence system. Case studies show that the framework is effective in automating geospatial analysis, improving the accuracy and completeness of knowledge extraction, and enhancing decision-making efficiency, while the hallucination and reliability of LLM remain open challenges.

Keywords

Large language model, Geographic information system, Geospatial knowledge graph, Spatial intelligence, Decision support, Self-evolution

Introduction

The volume of geospatial data is expanding at an unprecedented pace. Satellite remote sensing, unmanned aerial vehicle (UAV) surveys, ground sensor networks, and crowdsourcing platforms continuously generate spatiotemporal data. These data are multi-source, heterogeneous and multi-scale.

As a result, GIS has become a core decision-making tool. It supports applications in urban planning, environmental protection, disaster response, public health, and natural resource management. Traditional GIS, however, has long carried a high professional barrier: Spatial analysis requires users either to write code directly or to invoke tools indirectly through graphical interfaces. Workflows are hard to reuse, and ordinary decision makers struggle to turn complex spatial problems into executable analysis plans.

The emergence of large language models offers an opportunity to break this barrier. LLM are strong in semantic understanding, logical reasoning, and code generation. They can parse a user's natural-language intent into structured analysis tasks and then plan and invoke specialized tools autonomously. LLM4GIS systems such as Geoscience Generative Pre-trained Transformer (GeoGPT) and GIS Copilot have shown that natural-language-driven spatial analysis is feasible, giving rise to a new paradigm of autonomous GIS [1-3].

Existing research still has two notable shortcomings. Most LLM-driven GIS systems are task-execution static agents: They can complete a single analysis on command but lack the ability to discover and accumulate knowledge continuously, and the construction of geospatial knowledge graphs still relies heavily on manual annotation. More importantly, these systems generally do not self-evolve or improve with use. Task results, user feedback, and error-correction experience cannot be distilled into reusable knowledge, so system capability rarely rises over time.

To address these issues, this paper proposes a self-evolving spatial intelligence framework that deeply integrates LLM with GIS, forming a closed loop of knowledge discovery, decision support, and self-evolution. The framework first establishes a three-layer architecture consisting of a knowledge layer, an execution layer, and a decision layer, together with a self-evolution loop running through all layers.

This forms the methodological backbone of the paper. Building on this framework, the work proposes a reflexion-enhanced approach for automatic geospatial knowledge graph construction. In this method, a generator LLM and a critic LLM operate in a self-learning loop. Their collaboration improves the accuracy and completeness of knowledge extraction. The autonomous spatial analysis agent and the knowledge-augmented decision-support layer constitute the concrete form of the framework. Case studies are adopted to test its performance in spatial analysis automation, knowledge discovery, and decision efficiency. The remainder of this paper is organized as follows. First, it reviews related work; second, it describes the method; third, it reports the case studies and experimental results; fourth, it discusses open issues, and finally, it concludes the paper and points to future directions.

Related work

LLM and spatial analysis

The combination of LLM and spatial analysis is among the most active directions in Geo artificial intelligence (AI). GeoGPT is an early representative. Its design follows the paradigm of a foundation model coupled with specialized tools. The LLM acts as the reasoning core to schedule GIS tools following a think-plan-execute sequence.

This framework supports end-to-end workflows including data crawling, spatial analysis, and mapping. Systematic validation is conducted across scenarios of facility sitting, spatial query, and mapping.

Pisetta et al. conducted a review of 54 papers about LLM-based cartography and GIS. It identified four application clusters. These clusters include geospatial knowledge graph construction, multimodal data acquisition, language-driven spatial analysis and geometric question answering, and automated map styling. Meanwhile, they pointed out two trends. First, GPT-3.5/4.0 serves as the foundation for roughly two-thirds of the studies. Second, open-weight models such as LLaMA, Gemini, and DeepSeek are developing rapidly [4]. Progress is also visible in vision-language directions: MapReader learns a visual language model for map analysis, CartoAgent uses multimodal multi-agent collaboration for map style transfer and evaluation [5,6]. Feng et al. evaluated and enhanced the spatial cognition of LLM from a cognitive perspective [7]. These studies demonstrate the potential of LLM for spatial analysis, but most target a single task and have not yet formed a unified framework covering the whole chain of knowledge discovery, analysis, and decision-making.

Autonomous GIS and agents

Autonomous GIS aims to let a system complete spatial analysis independently with minimal human intervention. GIS copilot embeds an LLM into the quantum geographic information system (QGIS) platform, using an informed agent equipped with complete tool documentation to generate spatial analysis workflows automatically, and was evaluated on more than 100 tasks of varying complexity. Completion rates are high on basic and intermediate tasks, while fully autonomous execution of advanced tasks remains difficult. Data acquisition is another bottleneck, which Ning et al. addressed by proposing an autonomous GIS agent framework for data retrieval that pairs an LLM-driven decision core with a data-source handbook mechanism to discover and fetch multi-source spatial data. These efforts lay the engineering foundation for autonomous GIS, yet their position is still that of execution-oriented agents - completing tasks on demand rather than evolving systems that improve with use.

Geospatial knowledge graphs and knowledge discovery

A knowledge graph is a core carrier for the structured representation and reuse of spatial knowledge. LLM show few-shot and even zero-shot generalization in entity recognition and relation extraction, offering a new paradigm for automatically constructing geospatial knowledge graphs. The accuracy and completeness of such automatic construction. How to improve extraction quality in a controllable way is a central challenge in geospatial knowledge discovery and one of the problems this paper addresses.

Spatial reasoning and decision support

The ultimate purpose of spatial analysis is decision-making. Feng and Cao use a conversational agent to guide regionalization analysis, confirming the feasibility of LLM in lowering the barrier to spatial decision-making [8]. Yang et al. show that spatial cognition can be enhanced through targeted methods, supporting more reliable spatial reasoning and decision-making. Existing decision-support systems mostly rely on static knowledge bases and lack real-time linkage with knowledge discovery, making it hard to keep pace with new data and new tasks.

Research gap

Taken together, knowledge discovery, spatial analysis, and decision support are largely treated as separate stages in existing research, and a closed-loop mechanism that lets a system self-evolve with use is missing. This paper targets exactly this gap by proposing a unified framework in which a self-evolution loop connects the three stages.

Method

Overall framework

The self-evolving spatial intelligence framework consists of three layers, with a self-evolution loop running through them to form a closed loop. The knowledge layer centers on a geospatial knowledge graph. It extracts, organizes, and continuously updates spatial knowledge from multi-source heterogeneous data, providing structured semantic support for the upper layers [9].

The execution layer is an autonomous spatial analysis agent that parses natural-language intents, converts them into executable geospatial workflows, and completes the analysis through tool calling and multi-agent collaboration. The decision layer is a knowledge-augmented decision-support layer. It uses retrieval-augmented generation (RAG) to combine the knowledge graph with real-time analysis results and produces interpretable, evidence-based recommendations.

The self-evolution loop gathers task results, user feedback, and error-correction records. It feeds these signals back to the knowledge and execution layers. This mechanism supports incremental knowledge-based updates. It also enables continuous refinement of prompts and tool descriptions. As a result, system performance improves steadily during operation.

LLM-driven geospatial knowledge discovery module

This module aims to construct a geospatial knowledge graph automatically. Data ingestion is the first step, in which multi-source corpora such as geographic literature, technical reports, spatial database metadata, and open text are uniformly cleaned and normalized. The following entity recognition and relation extraction stage designs geography-oriented prompt templates, combined with few-shot examples, to guide the LLM in outputting subject-relation-object triples returned in structured JavaScript object notation (JSON). A single extraction pass tends to produce hallucinations and semantic ambiguity. For this reason, a reflexion-enhanced self-learning loop is adopted. In this loop, a generator LLM works together with a critic LLM. The generator completes the initial knowledge extraction. The critic carries out semantic verification and consistency checks. It identifies missed, incorrect and misaligned relations. Then the critic feeds its feedback to the generator. The generator repeats revisions iteratively until the output meets the predefined quality threshold. The verified triples are then written into a graph database and aligned with the existing knowledge graph through entity alignment and conflict resolution, enabling continuous knowledge accumulation [10].

Autonomous spatial analysis agent

The core of the execution layer is an autonomous spatial analysis agent that operates under the Reasoning and Acting (ReAct) paradigm. It first parses the user's natural-language intent and extracts spatial objects, analysis goals, and constraint parameters. It then selects and invokes appropriate tools from a tool library covering data loading, spatial analysis (buffering, overlay, clipping, interpolation, etc.), remote sensing, and mapping. The execution result is fed back to the agent as an observation, driving it to plan the next action until the task is complete. Complex tasks demand more collaboration, so a multi-agent mechanism is introduced in which a planning agent, an execution agent, and a verification agent divide the work, reducing the accumulation of a single LLM over long chains of tasks.

Knowledge-augmented decision support layer

The decision layer centers on retrieval-augmented generation. User queries are fused with the geospatial knowledge graph of the knowledge layer and the real-time analysis results of the execution layer, providing evidence context for the LLM, from which interpretable decision recommendations are generated. Compared with direct generation, the output of this layer carries reasoning chains and evidence sources, markedly reducing hallucination risk, and can be applied to facility sitting, urban planning, and disaster response to give evidence-based recommendations.

Self-evolution mechanism

Self-evolution is what distinguishes this framework from existing work. The system continuously collects three kinds of feedback signals: task execution results (success or failure), explicit user feedback (acceptance or correction), and error-correction records (tool-calling errors, extraction errors). After cleaning, the signals flow back through two pathways. Some signals incrementally update the knowledge graph and distill new spatial knowledge. Others refine tool descriptions and prompt templates.

This refinement corrects the LLM's misunderstanding of domain-specific terms. A typical example is clipping in GIS, whose meaning differs from the common everyday definition. Through this cycle, the system keeps correcting and strengthening itself during use, forming a virtuous loop of use, feedback, evolution, and improvement.

Case studies and experiments

Note: The numerical values presented in this section are for illustration, to demonstrate the expected performance and reasoning logic of the framework. These placeholder values should be replaced with real experimental data prior to formal submission.

Experimental setup

The knowledge-discovery corpus consists of several hundred texts drawn from geographic literature, technical reports, and spatial database metadata. The spatial analysis task covers three complexity levels: basic (single tool), intermediate (multi-step with guidance), and advanced (multi-step without guidance). Three kinds of baselines are used: direct LLM extraction (without the reflexion mechanism), traditional rule- or deep-learning-based extraction, and a single-agent (without multi-agent collaboration) spatial analysis scheme. Evaluation metrics cover precision, recall, and for knowledge extraction, task success rate for spatial analysis, and human-evaluation scores for decision support.

Knowledge discovery performance

In geographic entity recognition, the reflexion-enhanced mechanism rises by about 5 percentage points over direct LLM extraction; the improvement in relation extraction is even clearer, because the critic corrects misaligned spatial relations. Illustrative results give an entity-recognition of about 0.90 and a relation-extraction of about 0.82, both clearly better than traditional rule-based methods, with a large reduction in dependence on manual annotation.

Spatial analysis automation

The autonomous spatial analysis agent was tested on the spatial analysis task set. It reaches a completion rate of nearly 95% for basic tasks, around 80% for intermediate tasks, and approximately 60% for advanced tasks. This pattern aligns with the trend reported by Akinboyewa et al. Multi-agent collaboration further lifts the success rate of advanced tasks by 10 percentage points. This result shows that the separation of planning, execution, and verification effectively restrains error accumulation in long task chains.

Decision support effectiveness

In decision scenarios such as facility sitting and urban planning, the recommendations of the knowledge-augmented decision-support layer carry reasoning chains and evidence sources, and their human-rated interpret-ability and credibility both exceed those of direct generation. In terms of response efficiency, retrieval-augmented generation fused with the knowledge graph shortens multi-turn interaction time by about 30% on average.

Discussion

Despite its promise, the framework still faces several challenges. Hallucination and reliability come first: LLM still make mistakes when handling specialized geographic terms and complex spatial relations. Reflexion and knowledge augmentation can only mitigate these errors rather than eliminate them. Therefore, whether output remains consistent in professional scenarios is still an open question. Interpretability and trust are also critical concerns. The evidence chain of decision recommendations improves transparency. However, the internal reasoning of the model is hard to audit completely. On data quality and ethics, bias, privacy, and provenance issues in spatial data may be amplified by the model, so provenance tracing and ethical review mechanisms need to be built into the framework. Computational cost is another constraint: multi-agent collaboration and reflexion iterations incur higher overhead, and a trade-off between effectiveness and efficiency is required.

Conclusion

Existing LLM-driven GIS systems suffer from insufficient knowledge accumulation and autonomous evolution. To tackle this limitation, this paper proposes a self-evolving spatial intelligence framework. The framework integrates geospatial knowledge discovery and decision support. It adopts a three-layer architecture consisting of knowledge, execution, and decision layers, with a self-evolution loop running through all layers. Case studies preliminarily validate the framework's performance in knowledge-extraction accuracy, spatial-analysis automation, and decision-support effectiveness. Future work will focus on multimodal spatiotemporal data integration, causal spatial reasoning, and the construction of standardized evaluation benchmarks, moving GIS further from a passive tool toward a continuously evolving spatial intelligence system.

Funding

This work was not supported by any funds.

Acknowledgments

The author would like to show sincere thanks to those technicians who have contributed to this research.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Akinboyewa, T., Li, Z., Ning, H., Lessani, M. N. (2025) GIS copilot: towards an autonomous GIS agent for spatial analysis. *International Journal of Digital Earth*, 18(1), 2497489.

[2] Zhang, Y., Wei, C., He, Z., Yu, W. (2024) GeoGPT: an assistant for understanding and processing geospatial tasks. *International Journal of Applied Earth Observation and Geoinformation*, 131, 103976.

[3] Ning, H., Li, Z., Akinboyewa, T., Lessani, M. N. (2025) An autonomous GIS agent framework for geospatial data retrieval. *International Journal of Digital Earth*, 18(1), 2458688.

[4] Pisetta, J., Souza, F. A., Amorim, J., Antonio, N. D., Nunes, D. M., Araki, H., Camboim, S. P. (2026) Mapping with words: Integrating large language models into geospatial practice. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10, 285-292.

[5] Wang, C., Kang, Y., Gong, Z., Zhao, P., Feng, Y., Zhang, W., Li, G. (2025) CartoAgent: a multimodal large language model-powered multi-agent cartographic framework for map style transfer and evaluation. *International Journal of Geographical Information Science*, 39(9), 1904-1937.

[6] Zhang, Y., Zhang, W., Zeng, Z., Jiang, K., Li, J., Min, W., Luo, W., Guan, Q., Lin, J., Yu, W. (2025) MapReader: a framework for learning a visual language model for map analysis. *International Journal of Geographical Information Science*, 39(9), 1973-2008.

[7] Yang, A., Cheng, F., Jia, Q., Dong, W., Ma, M., Chen, H., Yang, F., Wu, H. (2025) Evaluating and enhancing spatial cognition abilities of large language models. *International Journal of Geographical Information Science*, 39(9), 2009-2044.

[8] Feng, X., Cao, Y. (2025) Enhancing the accessibility of regionalization techniques through large language models: a case study in conversational agent guidance. *International Journal of Geographical Information Science*, 39(9), 2045-2068.

[9] Zheng, W., Yang, M., Ren, Y., Wang, H., Zeng, C., Zhang, Y. (2026) CARAG: context-aware retrieval-augmented generation for railway operation and maintenance question answering over spatial knowledge graph. *ISPRS International Journal of Geo-Information*, 15(2), 78.

[10] Zhang, J., Qi, Y., Che, Y., Yang, R., Guo, X. K., Zhang, H. (2026) Resolving semantic conflicts in multi-source threat intelligence through temporal consistency modeling. *Information Sciences*, 123267.