

Mapping with Words: Integrating Large Language Models into Geospatial Practice

Jaqueline Pisetta, Fabíola A. Souza, Jaqueline Amorim, Nathan D. Antonio, Darlan M. Nunes, Hideo Araki, Silvana P. Camboim

Universidade Federal do Paraná (UFPR) - jaquelinepisetta@gmail.com; fabiola.andrade@ufba.br; jaqueline.urb15@gmail.com; nathandamas@ufpr.br; darlan.nunes@ufv.br; haraki@ufpr.br; silvanacamboim@gmail.com

Keywords: Cartography, ethics, semantics, language models, natural language processing.

Abstract

For decades, the focus of geospatial artificial intelligence (AI) has been on imagery and data, with linguistic interfaces remaining unexplored. This study presents a systematic review of research applying Large Language Models (LLMs) to Cartography and GIScience. We analysed 54 peer-reviewed articles published between 2023 and 2025, mapping the use of LLMs in data acquisition, semantic enrichment, spatial analysis, cartographic design, and user interaction, among other topics. The corpus reveals four dominant application clusters: (1) semantic creation/alignment of Geo-knowledge graphs; (2) text- and vision-based data acquisition; (3) language-driven spatial analysis and Geometric Question Answering (GeoQA); and (4) automated map styling and symbolisation. GPT-3.5/4 underpins two-thirds of the studies, while open-weight models, such as LLaMA-2, FLAN-T5, Gemini, and DeepSeek, are gaining traction. LLM work aligns most strongly with the challenges of openness and reproducibility, as well as cartographic design automation, but is noticeably weaker in areas such as provenance ethics, causal inference, participatory mapping, and mobile multimodal interaction. We outline three priorities for future research: (i) open benchmark datasets for spatial reasoning and map quality; (ii) ethics checklists that surface bias, privacy and hallucination risks; and (iii) investment in multilingual, low-resource Geo-LLMs to broaden global participation. By mapping current advances against long-standing research gaps, the review provides an actionable agenda for guiding large language models (LLMs) toward equitable and trustworthy cartographic practice.

1. Introduction

1.1 Why Language Matters to Mapping

People have always interacted with geographical spaces through language. Throughout history, communities have named mountains, rivers and settlements, creating a common reference system that connects stories, rights and memories to particular places. Research into place names (toponymy) shows that these labels are not merely passive entries in a directory but an active socio-cultural practice. Equally important is the way language has been used to describe the relationships between real-world objects in space. Expressions such as 'beyond the forest', 'south of the river' or 'behind the temple' organise everyday navigation and narrative. Together, naming and relational description form the oldest and most pervasive human 'interface' with geographic space.

Despite the language being our original spatial interface, the first three decades of Geospatial Artificial Intelligence (GeoAI) paid comparatively little attention to it. For three decades, the majority of GeoAI efforts have concentrated on remote sensing and photogrammetry, with the development of machine learning and various neural network models enabling a broad range of image-based applications (Chamma et al., 2021). Notably, even in countries with limited financial resources for scientific research, such as Brazil, a significant volume of work in this area has been produced (Bastos et al., 2023; Silva et al., 2023; Moreira et al., 2023). Comprehensive reviews of deep-learning change-detection methods demonstrate the maturity and scale of this subfield.

Natural language processing (NLP) techniques have had limited applications in Geographic Information Science (GIScience) for over two decades, including gazetteer matching and geoparsing. However, the recent surge in transformer-based large language models represents a significant change.

Although NLP has established a presence in the field, the use of transformer-based large language models is only now gaining momentum, with most papers dating back to 2022 or later. This temporal lag explains why critical issues of bias, provenance, and

evaluation remain underexplored in the language domain despite image-based GeoAI approaches becoming routine. This imbalance is addressed in the remainder of the study, which explores where and how Large Language Models (LLMs) are being incorporated into Cartographic practice, the challenges that remain and how future work can integrate linguistic, visual and numerical intelligence into a coherent geospatial science.

1.2 Large Language Models across the Mapping workflow

The interpretation of human language within the framework of machine learning is a particularly notable application of AI, as NLP continues to expand its potential applications. In NLP, algorithms identify patterns and generate responses through linguistic modelling. Traditional approaches involve interpreting text sequences syntactically and semantically, applying fundamental mathematical models to existing training datasets and deriving new classifications (Mai et al., 2022; Jozefowicz et al., 2016). More recent approaches are based on neural networks, with the Transformer architecture, introduced by Vaswani et al. (2017), revolutionising text analysis by mitigating the challenges of handling long-range dependencies and reducing processing time through parallelisation. This architecture has been instrumental in the development of LLMs, which can process vast amounts of data (Q. Zhang et al., 2024; Strafforello, 2023). Because language is so fundamental, it surfaces at three distinct points in the geospatial workflow.

- **Data acquisition:** Modern NLP techniques now parse free-text addresses, social media posts, and street-level imagery, converting everyday expressions into structured features.
- **Human-data interaction:** Voice and chat-based assistants. The shift from SQL and menu clicks to conversational interfaces adds inclusivity (multilingual support) and accessibility (hands-free or low-vision use).
- **Cognitive response:** Spatial-language studies show that the words we use to encode direction, distance and topology influence the mental maps we build and how confidently we navigate. Research on cognitive maps and children's spatial-

language development suggests that richer linguistic repertoires are associated with stronger spatial memory and analytic skills. Consequently, legends, captions and voice guidance are not mere decorations - they shape users' interpretation of spatial patterns.

1.3 Persistent research challenges in Cartography & GIScience

Recent agenda papers by Çöltekin et al. (2017), Griffin et al. (2017), Robinson et al (2017), Kang et al. (2024), Fish et al. (2025), and Gartner et al. (2025), as well as by Nelson et al. (2025) - show that many of the problems initially identified by Bertin (1983), MacEachren (1995), MacEachren and Kraak (2001), as well as the ICA's research roadmaps from the 2000s, remain unresolved. We therefore summarise these long-standing concerns in Table 1, grouping them into eight broad clusters that set the stage for our review.

Challenge cluster	Persistent issue
1 - Human factors and cognition	<ul style="list-style-type: none"> Understanding how abilities, disabilities, emotions, and context shape map reading and wayfinding and mitigating over-reliance on turn-by-turn navigation.
2 - Data ethics and provenance	<ul style="list-style-type: none"> Ensuring fair representation, avoiding selection bias, safeguarding privacy, and documenting lineage and accountability in map production.
3 - Openness & reproducibility	<ul style="list-style-type: none"> Sharing data, code and design decisions so that cartographic research and products can be verified and reused.
4 - Scale, big data and uncertainty	<ul style="list-style-type: none"> Abstracting, generalising and visualising massive, heterogeneous and sometimes low-quality geospatial datasets while conveying uncertainty.
5 - Cartographic design automation	<ul style="list-style-type: none"> Achieving high-quality symbolisation, typography and layout without intensive manual effort or loss of artistic nuance.
6 - Causality & spatial inference	<ul style="list-style-type: none"> Moving beyond descriptive patterns to establish causal explanations despite spatial autocorrelation and confounding.
7 - Mobile and multimodal interaction	<ul style="list-style-type: none"> Delivering clear, effective maps on small, sensor-rich devices and through mixed modalities (touch, voice, haptics).
8 - Collaboration and participatory mapping	<ul style="list-style-type: none"> Enabling non-experts to co-create, critique and update maps while managing conflicts and quality control.

Table 1. Research Challenges in Cartography & GIScience.

1.4 Aim and scope of this contribution

This contribution presents a state-of-the-art synthesis of how researchers are deploying LLMs in cartography and GIScience, identifies common use cases and models, maps global research distribution, highlights technical and ethical gaps, and issues recommendations for future inquiry.

2. Materials and methods

2.1 Search strategy, Screening & Inclusion

To map scientific production at the intersection of LLMs and Cartography, groups of search terms were defined based on key themes. These groups guided the formulation of queries in Scopus and were organised according to some dimensions: technology (LLM/GeoAI), cartographic domain, and semantic and ethical aspects. Combining these groups made it possible to identify trends in areas such as geospatial data acquisition, map design, toponymy, spatial narratives, and ethics in automated cartography. Table 2 summarize the blocks used in the queries.

Grouped Terms ("Large Language Model" AND)
<ul style="list-style-type: none"> geoai; gis; geospatial; geospatial systems; geospatial data; mapping; cartography; map design; symbolization; and maps; cartographic design; storymaps; cognitive maps; geoai; and cartography; and large language model; geoai; and cartographic design; geoai; and cartographic visualization; geoai; and cartographic generalization; semantic; ontology; knowledge-based systems; knowledge-based systems; and geospatial; ontology; and geospatial; knowledge-based systems; and ontology; natural language; and geospatial data; natural language; and maps; and cartography; toponym; toponym; or toponyms; or place names; or geographic names; gazetteer; and cartography; geoai; and openstreetmap; openstreetmap; ethics; and cartography; ethical geospatial data acquisition; ai ethics; and cartography; geoespatial data acquisition; semantic analysis; andgeospatial; geospatial storytelling; requirements engineering; and cartography; design thinking; and cartography; agile methodologies; and cartography; cartographic knowledge; geoai; and cartographic data;

Table 2. Terms used in queries on the Scopus database.

From the search on the Scopus database, using the Boolean logic in Table 2, articles, abstracts, and short papers were filtered, prepared in English or Portuguese, with open access, and without a time filter.

2.2 Data extraction

After systematically selecting scientific documents, we began a detailed data extraction process, which was fundamental for synthesising information relevant to this research's objectives. The rigorous application of filters and query groups outlined in the "Search strategy, Screening & Inclusion" section allowed us to identify a manageable, highly focused corpus. As a result of this careful filtering, 54 articles that directly align with the intersection of LLMs and the geospatial and cartographic domains were selected. Articles were excluded if they were not peer-reviewed, did not present a direct application of LLMs to Cartography or GIScience, duplicated content already retrieved from other sources. These articles constitute the empirical basis for this review and demonstrate an increasing volume between 2023 and May 2025 (Figure 1).

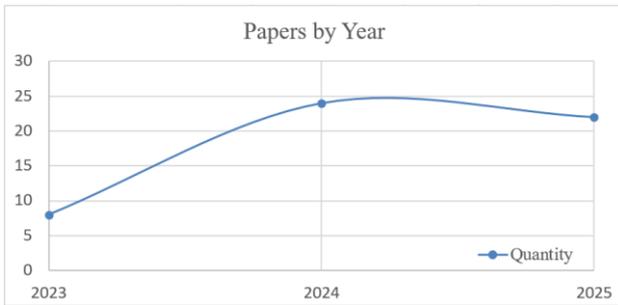


Figure 1. Papers by year.

The fields extracted for analysis include year of publication, title of the paper, authors, institution and country of the main author, journal, type of access, language, and abstract. The objective, the LLM used and the stage in the workflow were also identified. All the papers are listed in Appendix A (<https://anonymous.4open.science/r/MappingWords-834E>).

3. Findings

3.1 LLMs for describing the world: concepts, ontologies, semantics

Large language model research on geospatial semantics now spans the entire knowledge graph lifecycle. The ten papers reviewed here illustrate four successive stages of development. First, LLMs create or extract initial triples directly from raw corpora. Second, they align or extend heterogeneous ontologies, turning isolated datasets into interoperable Geospatial Knowledge Graphs (GeoKG). Third, they enable reasoning and Geometric Question Answering (GeoQA), answering conversational queries about topological and directional relations without SQL. Finally, they support assessment and governance, auditing the completeness and consistency of GeoKG. Table 3 maps each study to its dominant stage, revealing where effort is concentrated and where opportunities remain.

Together, these groups are pushing GIScience towards semantically aware GeoAI pipelines, where text, code, and spatial relationships are all grounded in a shared knowledge base.

Cluster	n	Papers
Creation / Extraction	3	(He et al., 2024; Hou et al., 2025; Ji and Gao, 2023)
Alignment / Extension	3	(Liang et al., 2025; Wu et al., 2024; Tang et al., 2023)
Reasoning / GeoQA	2	(Kazemi et al., 2025; Yan and Lee, 2024)
Assessment / Governance	2	(Souza et al., 2025; S. Wang et al., 2024)

Table 3. LLMs across the geospatial knowledge graph lifecycle.

3.2 LLMs for data acquisition

LLMs are also transforming the way data is acquired in GIScience. The twenty-two studies collected for this cluster demonstrate how LLMs transform unstructured content—tweets, news stories, addresses, entire corpora, and even raw imagery—into georeferenced datasets that can be directly fed into spatial analysis pipelines. Table 4 organises the literature into four functional areas:

- **Toponyms and Geocoding:** These studies utilise LLMs to recognise place names, historical toponyms, or fuzzy

address strings, resolve ambiguities against gazetteers, and return precise coordinates.

- **Vision–Language / Imagery Extraction:** Multimodal models link textual prompts to objects in street-level photos, remote-sensing scenes, or OCR’d map sheets, producing automatically geotagged footprints or labels. This line of work suggests end-to-end pipelines where imagery, captions, and map layers are generated with minimal human intervention.
- **Social-Media / News Event Tagging:** Here, LLMs detect hazards, crises, or points of interest mentioned in real-time feeds - such as tweets, Telegram posts, and breaking news wires - and attach spatial footprints. The aim is rapid situational awareness for emergency response or live dashboards.
- **Other Data-Extraction Tasks:** These cover specialised areas, including building corpora of Indigenous place descriptions, mining legal documents for land parcel references, and synthesising benchmark datasets for future large language model (LLM) evaluation. Collectively, they broaden the data spectrum beyond the mainstream text-and-image sources.

Cluster	n	Papers
Toponyms and Geocoding	11	(Chen et al., 2025; S. Li et al., 2025; H. Liu and Mao, 2025; Sui et al., 2025; Windiastuti et al., 2025; Gomes et al., 2024; Hu and Kersten, 2024; Hu et al., 2024; Huang et al., 2024; Kim and Lee, 2024; Kenyon et al., 2023)
Vision-Language / Imagery	4	(Han et al., 2025; Hsu et al., 2024; W. Li et al., 2024; Y. Zhuang et al., 2024)
Social media / News	3	(Tian et al., 2025; Yin et al., 2025; Klyachin and Khizhnyakova, 2024)
Specialised domain extraction	4	(Y. Liu et al., 2025; Y. Li et al., 2024; Lipka et al., 2024; Z. Li and Zhang, 2023)

Table 4. Data Acquisition Group.

Together, these approaches are pushing GIScience towards LLM-enabled data pipelines, in which text, images, and crowd contributions are transformed into spatial layers with significantly less manual intervention.

3.3 LLMs for spatial analysis and interactions

LLMs are now venturing from data preparation into spatial analysis itself, helping users pose questions, uncover patterns and automate modelling steps that once required specialised GIS scripting. The twelve papers in Table 5 illustrate distinct ways LLMs are being woven into analytic workflows:

- **Interaction & NL interfaces** papers focus on how users ask or receive information (natural-language queries, chatbots, GeoQA).
- **Spatial analysis and geoprocessing** papers focus on the analyses the system performs (buffering, classification, disaster assessment), often generating Python scripts behind the scenes.

Cluster	n	Papers
Interaction & NL interfaces	5	(Jeong et al., 2025; Xie et al., 2025; Xu and Tao, 2024; Feng et al., 2023; Tao and Xu, 2023)
Spatial analysis / geoprocessing	7	(L. Xu et al., 2025; Mansourian and Oucheikh, 2024; Morandini et al., 2024; Nie and Gao, 2024; Q. Zhang and Gao, 2024; Y. Zhang et al., 2024; Du et al., 2023)

Table 5. Spatial Analysis Group.

Together, these studies demonstrate an increasing shift in GIScience toward language-driven spatial analysis. In this approach, users can initiate and interpret complex spatial operations using natural language. However, establishing trust and accuracy through robust quality assessment frameworks remains a challenge.

3.4 LLMs for map symbolisation and design

LLMs - and, increasingly, diffusion-based text-to-image models - are beginning to automate classical map design chores, from symbol and palette suggestions to fully generative styling. The five papers in Table 6 fall into two practical clusters:

- Automated map styling and template: Models propose complete visual styles or apply diffusion methods (text-to-image) to generate map products with minimal manual intervention.
- Label placement and typography: LLM-guided heuristics optimise label positions or font choices to enhance readability while respecting cartographic conventions.

Cluster	n	Papers
Automated map styling and template	3	(Memduhoğlu, 2025; Dunkel et al., 2024; C. Wang et al., 2024)
Label placement and typography	2	(C. Wang et al., 2025; Y. Zhang et al., 2024a)

Table 6. Cartographic Design and Symbolisation Group.

Together, these studies demonstrate how large language models can support aesthetic reasoning and expressive design decisions that were previously considered too subjective and manual for computers to handle. This points toward a future of intelligent, inclusive, and collaborative mapmaking.

Beyond these two clusters, GeoAI literature shows a broader spectrum of cartographic design automation that can deepen this discussion. Kang and Liu. (2022) review how deep generative models are applied not only for map style transfer but also for evaluating colour schemes and visual quality, automating aspects of readability and aesthetics that traditionally relied on expert judgement. This work suggests that large models can move beyond symbol placement and palette suggestion to support systematic design evaluation, keeping humans “in the loop” to refine outputs while benefiting from machine-driven insights. Such advances indicate a research trajectory toward an “artificial cartographer assistant” capable of complementing expert design reasoning with scalable, data-driven feedback.

3.5 Cross-cutting patterns

A clear publication pattern emerges across the 54-item corpus (38 journal articles and 16 conference papers — see Figure 2), retrieved primarily through a structured search in the Scopus database. The journal papers cluster around two key publications: the ISPRS International Journal of Geo-Information and the

International Journal of Geographical Information Science (IJGIS). Together, these two outlets account for almost half of all peer-reviewed articles.

Conference contributions, meanwhile, are more geographically dispersed. However, they still converge on a small set of venues, including recent GIScience and GeoAI meetings in Glasgow, Athens, New York, Leeds, Los Angeles, and Bangkok. The author-affiliation metadata (see Figure 3) reveals a significant concentration in the United States and China, with only a few contributions from continental Europe. Taken together, these patterns suggest that, although LLM-focused GeoAI research is finding broader publication channels, it remains linguistically and geographically imbalanced — an issue that future work will need to address.

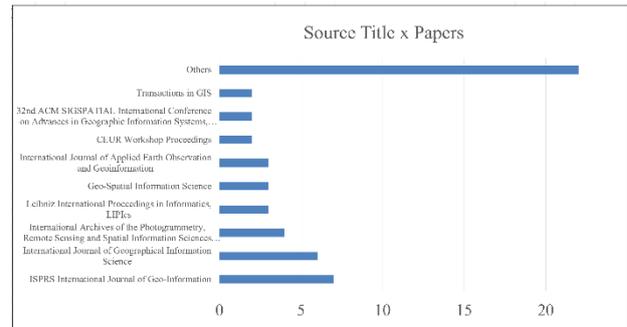


Figure 2. Publication venues.

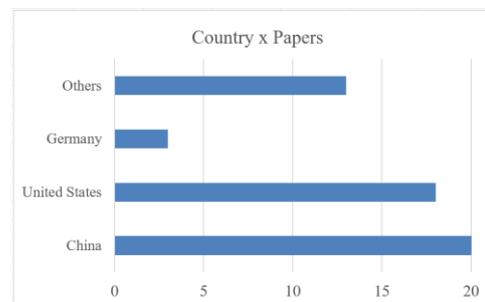


Figure 3. Geographical distribution.

Although OpenAI's GPT series continues to dominate the corpus — with ChatGPT/GPT-3.5-turbo and GPT-4 featuring in around two-thirds of papers — the landscape is gradually diversifying (Figure 4). Around a fifth of the studies use open-source models such as LLaMA-2, FLAN-T5 or variants of the BERT family for fine-tuning and reproducibility purposes. A smaller but growing number of experiments with newly released foundation models: For example, Gemini (Google) is tested in two papers for its ability to understand multimodal prompts, pairing textual instructions with street-level imagery, while DeepSeek-6B/DeepSeek-Coder is used in three disaster analysis or code generation studies. These models are valued for their Chinese-English pre-training mix and open license. Finally, a small amount of work utilises vision-language models, such as CLIP or BLIP-2, for image-based place extraction. Almost all projects adapt these foundations using prompt engineering or lightweight fine-tuning rather than training from scratch, which highlights the current GeoAI focus on repurposing general LLMs for spatial tasks rather than developing specialised Geo-LLMs from scratch.

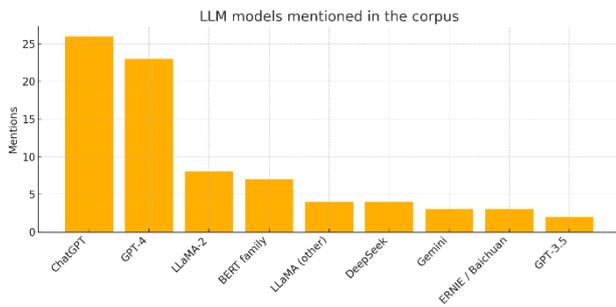


Figure 4. Frequency of LLMs applied.

4. Discussion

Integrating LLMs into cartography and GIS represents a significant shift in how spatial data is processed, defined, and communicated. This discussion explores how LLMs can enhance technical workflows, including geocoding, spatial querying, and symbolisation. LLMs can also transform the human-machine interface by providing more inclusive, conversational, and semantically rich tools. While these advances are promising, some gaps require further research related to evaluation, ethical oversight, and spatial accuracy in the application of Artificial Intelligence (AI) to Cartography and GIScience.

4.1 Ethical implications

LLMs have made substantial advancements in Cartography. However, their adoption raises ethical concerns. A key issue is that LLMs can be trained using large and often uncurated datasets, which risk perpetuating socio-spatial inequalities. For example, they may underrepresent marginalised communities in geographic narratives or toponymy (Z. Liu et al., 2025; Q. Zhang et al., 2023). Another critical concern is provenance transparency, as noted by Dibiase (2025). Ethical cartographic practice requires the traceability of spatial information; however, LLM-generated content often obscures the origins of data, which undermines its reliability and accountability.

Additionally, the risk of LLMs producing geographically plausible yet fabricated descriptions can have serious consequences in sensitive areas such as disaster management and land governance (Romano, 2025). The unequal distribution of computational infrastructure exacerbates these issues, as access to high-performance large language models (LLMs) is often limited to well-funded institutions, thereby widening the global disparity in geospatial innovation (Mai et al., 2025).

These challenges reinforce the need for robust evaluation frameworks, transparency in model development and equitable investment in multilingual, context-sensitive AI capacity building, as emphasised in recent GeoAI literature and ethical frameworks.

4.2 Positive prospects

The potential for LLMs to transform cartography is significant, especially in terms of reducing the time spent on manual tasks. These models can automate processes such as symbol selection, layout refinement, and annotation, streamlining map production and enabling quicker prototyping and iterative improvements. In some cases, LLMs can also suggest suitable design models that can be refined as needed.

Their multilingual capabilities broaden accessibility further by enabling spatial queries and interactions in various languages, eliminating the need for separate translation tools. This makes

maps more adaptable and engaging for a broader audience. Additionally, LLMs offer a practical solution for small agencies and under-resourced institutions by replacing complex coding environments with natural language interfaces, enabling these organisations to carry out advanced spatial operations more efficiently.

These tools have the potential to improve users' perception and understanding of maps by supporting the creation of dynamic, personalised designs that adapt to individual need and context. Overall, these advancements suggest a more inclusive, responsive, and perceptually effective future for cartographic work.

4.3 Limitations of current literature

Despite the growing interest in applying LLMs to Cartography and GIScience, the existing literature reveals research gaps. Notably, there is a lack of benchmark datasets designed to evaluate geospatial reasoning and cartographic quality. Language coverage is also limited, with most research focusing on English and Chinese. This restricts the inclusivity and cross-cultural relevance of these tools. Furthermore, studies examining the cognitive impact of AI-generated cartographic outputs on users, including trust, interpretation, and interaction, are scarce. Ethical considerations are also understudied. These gaps suggest an urgent need for broader, multidisciplinary inquiry into how these tools are reshaping the field, both technically and socially.

Another critical but underexplored dimension is user trust and cognitive ergonomics. While LLM-driven tools can automate symbolisation and narrative generation, the interpretability of their outputs remains uncertain and user-dependent. Trust in AI-generated maps is shaped by cognitive load, visual clarity, and the user's own spatial literacy; yet few studies have systematically tested these aspects. Future work should include empirical user studies and cognitive evaluation to ensure that LLM-driven cartographic products are not only technically correct but also understandable and actionable by their intended audiences.

4.4 Alignment with persistent research challenges in Cartography and GIScience

Section 1.3 (see Table 1) outlines the eight persistent challenges that have shaped cartographic and GIScience research. To understand how recent literature centred on LLMs responds to these issues, we cross-referenced every paper in our review against the same set of challenges:

- Reproducibility and openness are still the main focus, with most of the corpus release code, benchmarks and open GeoKGs confirming a strong shift towards transparent GeoAI practice.
- Design and workflow automation are advancing – several studies employ prompt-based styling and analysis, GPT-guided symbology, and diffusion-derived map templates to address a long-standing bottleneck in cartographic production and map usage.
- Big-data scalability remains central, with papers coupling LLMs with streaming imagery or social media feeds. However, techniques for conveying uncertainty remain mostly ad hoc.
- Although human-factor research is growing, many papers are still shallow. While some introduce natural language or voice interfaces, only a handful conduct rigorous cognitive or accessibility evaluations, leaving room for deeper user-centred work.

- Ethics and provenance are less prevalent: fewer papers audit bias, privacy or lineage, indicating that the field is still evolving in line with broader AI governance debates.
- Causal inference is in its initial stages – only three studies attempt counterfactual or qualitative spatial reasoning, indicating a preliminary but promising area of research.
- Mobile and multimodal interaction are emerging, while collaborative and participatory mapping are scarcely addressed — both of which are fertile grounds for future LLM research in GIScience.

Overall, LLM-centric GeoAI is making tangible progress on several persistent challenges identified in earlier agendas—most notably openness and automated design—but considerable gaps remain in provenance, user cognition, causal reasoning and participatory workflows, where the next wave of studies can deliver the greatest impact.

5. Conclusions

This study aimed to systematically map the integration of LLMs into the domains of Cartography and GIScience, highlighting key application areas, identifying significant gaps, and exploring the ethical and technical challenges associated with this integration. Through a comprehensive literature search, screening, and analysis of 54 relevant peer-reviewed articles published between 2023 and May 2025, we provided insights into the current state of the art, revealing an emphasis on data acquisition, semantic enrichment, and interactive dissemination.

The findings indicate that LLMs have considerably influenced spatial data processing by automating traditionally manual tasks, enhancing semantic capabilities, and democratising cartographic knowledge through conversational interfaces. However, critical gaps remain, including the limited availability of benchmarks for spatial reasoning, insufficient development of multilingual models beyond English and Chinese, and persistent concerns regarding biases and ethical transparency.

To address these challenges and advance the responsible integration of LLMs into Cartography, we recommend the publication of open evaluation corpora to foster consistent, transparent, and reproducible evaluation of LLM performance in geospatial contexts. While a fully developed, widely accepted taxonomy of evaluation criteria would require additional empirical work and consensus-building beyond the scope of this review, the four application clusters identified here can serve as an initial conceptual scaffold to guide future efforts in designing such frameworks.

Additionally, comprehensive ethics checklists should be developed for GeoAI projects to proactively manage algorithmic biases, ensure transparency of provenance, and mitigate potential "hallucinations," thereby ensuring the responsible deployment of AI. Furthermore, investments in the development and training of non-English LLMs are essential to promote inclusive GIS practices and broaden global participation, especially for underrepresented linguistic communities.

We emphasise that the integration of the transformative potential of LLMs into cartographic workflow is neither inevitable nor deterministic. The future trajectory will be shaped by deliberate and thoughtful engagement from researchers, practitioners and policymakers to guide these technologies towards beneficial, equitable and ethically applications in cartography and GIScience.

References

Bastos, R.M.C., Machado, G.F. de S., Frauches, N., Cruz, C.B.M., 2023: Potencialidades e aplicações do uso de deep

learning no sensoriamento remoto em análises urbanas. In: *Simpósio Brasileiro de Sensoriamento Remoto*, São José dos Campos: INPE. Available at: <https://proceedings.science/sbsr-2023/trabalhos/potencialidades-e-aplicacoes-do-uso-de-deep-learning-no-sensoriamento-remoto-em?lang=pt-br>.

Bertin, J., 1983. *Semiology of graphics*. University of Wisconsin press.

Chamma, W.D.S., et al., 2021: Machine learning applied to satellite imagery for rooftop classification. *Brazilian Journal of Development*, 7(7), 72558–72576.

Çöltekin, A., Bleisch, S., Andrienko, G., Dykes, J., 2017. Persistent challenges in geovisualization – a community perspective. *International Journal of Cartography* 3, 115–139. <https://doi.org/10.1080/23729333.2017.1302910>.

DiBiase, D., 2025: Practical Geospatial Ethics: Concerns, Codes, and Cases. *Cartogr. Perspect.*, 105. <https://doi.org/10.14714/CP105.1935>.

Fish, C.S., Roth, Robert E., Raposo, Paulo, Guidero, Elaine M., and Sui, Z., 2025. Why conduct cartographic research: a comparison of cartographic research contributions in four International Cartographic Association (ICA) affiliated journals. *International Journal of Cartography* 11, 190–215. <https://doi.org/10.1080/23729333.2025.2475411>.

Gartner, G., 2025. A decade in scientific cartography: insights and future directions. *International Journal of Cartography*.

Griffin, A.L., White, T., Fish, C., Tomio, B., Huang, H., Sluter, C.R., Bravo, J.V.M., Fabrikant, S.I., Bleisch, S., Yamada, M., Picanço, P., 2017. Designing across map use contexts: a research agenda. *International Journal of Cartography* v. 1, 1–25. <https://doi.org/10.1080/23729333.2017.1316466>.

Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N., Wu, Y., 2016: Exploring the Limits of Language Modeling. *arXiv*, 11 February. <http://arxiv.org/abs/1602.02410>.

Kang, N. and Liu, C., 2022. Towards landscape visual quality evaluation: methodologies, technologies, and recommendations. *Ecological Indicators*, 142, p.109174. <https://doi.org/10.1016/j.ecolind.2022.109174>

Kang, Y., Gao, Song, and Roth, R.E., 2024. Artificial intelligence studies in cartography: a review and synthesis of methods, applications, and ethics. *Cartography and Geographic Information Science* 51, 599–630. <https://doi.org/10.1080/15230406.2023.2295943>.

Liu, Z., Janowicz, K., Majic, I., Shi, M., Fortacz, A., Karimi, M., Currier, K., Mai, G., 2025: Operationalizing Geographic Diversity for the Evaluation of AI Generated Content. *Trans. GIS*, in press. <https://doi.org/10.1111/tgis.14192>.

MacEachren, A.M., Kraak, M.J., 2001. Research challenges in geovisualization. *Cartography and geographic information science*, 28(1), pp.3-12. <https://doi.org/10.1559/152304001782173970>

MacEachren, A.M., 1995. *How Maps Work: Representation, Visualization, and Design*, 1st edition. ed. The Guilford Press, New York.

Mai, G., Cundy, C., Choi, K., Hu, Y., Lao, N., Ermon, S., 2022: Towards a foundation model for geospatial artificial intelligence (vision paper). In: *Proc. 30th Int. Conf. Advances in Geographic*

- Information Systems (SIGSPATIAL '22)*, 1–4. <https://doi.org/10.1145/3557915.3561043>.
- Mai, G., et al., 2024: On the Opportunities and Challenges of Foundation Models for GeoAI. *ACM Trans. Spatial Algorithms Syst.*, 10(2), 46. <https://doi.org/10.1145/3653070>.
- Mai, G., Xie, Y., Jia, X., Lao, N., Rao, J., Zhu, Q., Liu, Z., Chiang, Y., and Jiao, J. 2025. Towards the next generation of Geospatial Artificial Intelligence. *International Journal of Applied Earth Observation and Geoinformation* 136, 104368. <https://doi.org/10.1016/j.jag.2025.104368>
- Moreira, L., Blanco, L.C.M., Antonio, N.D., Araki, H., 2023: Utilização do algoritmo de Generative Adversarial Network em super resolução de imagens para o aumento de qualidade de modelos de elevação digitais baseados em dados SRTM. In: *Simpósio Brasileiro de Sensoriamento Remoto*, Florianópolis: INPE. <https://proceedings.science/sbsr-2023/papers/utilizacao-do-algoritmo-de-generative-adversarial-network-em-super-resolucao-de?lang=en>.
- Nelson, T., Frazier ,Amy E., Kedron ,Peter, Dodge ,Somayah, Zhao ,Bo, Goodchild ,Michael, Murray ,Alan, Battersby ,Sarah, Bennett ,Lauren, Blanford ,Justine I., Cabrera-Arnau ,Carmen, Claramunt ,Christophe, Franklin ,Rachel, Holler ,Joseph, Koylu ,Caglar, Lee ,Angela, Manson ,Steven, McKenzie ,Grant, Miller ,Harvey, Oshan ,Taylor, Rey ,Sergio, Rowe ,Francisco, Şalap-Ayça ,Seda, Shook ,Eric, Spielman ,Seth, Xu ,Wenfei, and Wilson, J., 2025. A research agenda for GIScience in a time of disruptions. *International Journal of Geographical Information Science* 39, 1–24. <https://doi.org/10.1080/13658816.2024.2405191>.
- Robinson, A.C., Demšar, U., Moore, A.B., Buckley, A., Jiang, B., Field, K., Kraak, M.-J., Camboim, S.P., Sluter, C.R., 2017. Geospatial big data and cartography: research challenges and opportunities for making maps that matter. *International Journal of Cartography* 3, 32–60. <https://doi.org/10.1080/23729333.2016.1278151>.
- Romano, A., 2025: Synthetic geospatial data and fake geography: A case study on the implications of AI-derived data in a data-intensive society. *ARPI – Univ. Pisa*, 1 January. <https://arpi.unipi.it/handle/11568/1279127>.
- Silva, B.L.C., Campanharo, W.A., Souza, F., Ferreira, K., Queiroz, G.R., 2023: Deep learning as a tool to interpolate cloudy pixels in Sentinel-2 time series. In: *Simpósio Brasileiro de Sensoriamento Remoto*, São José dos Campos: INPE. <https://data.inpe.br/bdc/web/en/deep-learning-as-a-tool-to-interpolate-cloudy-pixels-in-sentinel-2-time-series-2/>.
- Strafforello, F., 2023: How Transformers and Large Language Models (LLMs) Work — A Comprehensive Guide Using BERT, GPT, and T5. *Medium Magazine*, 11 June. <https://blog.gopenai.com/how-transformers-and-large-language-models-llms-work-3f20bb41c1ff>.
- Vaswani, A., Shazeer, N., Parmar, N., et al., 2017: Attention is All You Need. In: *Proc. 31st Conf. Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA.
- Zhang, Q., Zhang, T., Zhai, J., et al., 2024: A Critical Review of Large Language Model on Software Engineering: An Example from ChatGPT and Automated Program Repair. *arXiv*, 17 April. <http://arxiv.org/abs/2310.08879> (Accessed: 28 April 2024).
- Zhang, Q., Kang, Y., Roth, R., 2023. The Ethics of AI-Generated Maps: DALL·E 2 and AI's Implications for Cartography. Presented at the Leibniz International Proceedings in Informatics, LIPIcs. <https://doi.org/10.4230/LIPIcs.GIScience.2023.93>.