Journal of Global Research in Electronics and Communication
Volume 1, No. 6, June 2025
Available Online at: www.jgrec.com
RESEARCH PAPER

Analyzing the Performance of Large Language Models in Complex Spatial Reasoning Tasks

Rajiv Kumar
Director - Data Science
Oracle
USA
groverrajiv1984@gmail.com

***Abstract*—Spatial reasoning, which involves understanding the relationships between objects or entities in space, is a fundamental aspect of human cognition but remains a significant challenge for large language models (LLMs). This paper explores spatial reasoning, a critical but challenging task for large language models (LLMs) that requires understanding the relationships between spatial entities points, lines, and regions through metric, topological, directional, and order relations. A specialized evaluation dataset focused on Australian geography was developed to minimize pre-training bias, featuring 239 carefully crafted spatial reasoning questions. Fifteen prominent LLMs from OpenAI, Google, Anthropic, Meta, and Mistral were assessed under controlled zero-shot conditions across five key experiments: Toponym Resolution, Metric Relations, Directional Relations, Topological Relations, and Cyclic Order Relations. Results revealed significant variation in model performance, with models struggling notably in metric and cyclic order tasks, while showing relatively better outcomes in qualitative topological reasoning. Metric relation errors often involved underestimations of distance, directional reasoning accuracy declined with task complexity, and cyclic order reasoning approached random performance. Models from Google and Anthropic demonstrated greater caution, abstaining more frequently in the face of uncertainty, highlighting the ongoing challenges LLMs face in mastering complex spatial reasoning tasks.**

Keywords—Spatial Reasoning, geo-foundation, LLMs, geospatial, topological, Metric Relations, Embedding Techniques, self-supervised training.

I. INTRODUCTION

Every day, humans use spatial reasoning to interpret items in their surroundings. Navigation from one location to another, identifying places by adjacent landmarks, and avoiding collisions with other moving things would be difficult without significant world knowledge, experience, spatial intuition, common sense, and embodiment [1][2]. Objects' movement across space necessitates constantly acquiring data and deducing implicit knowledge, making spatial reasoning an especially difficult kind of thinking [3]. Using formal techniques applied to certain data formats has traditionally been the go-to solution for many spatial reasoning problems, such as spatial pattern matching (finding objects in the environment that fit a set of spatial restrictions) [4][5]. The types of issues that traditional techniques can tackle are limited by their processing slowness and rigidity, as well as their tendency to employ pre-computed indexes and data structures [6].

Recent studies have investigated the kind of world knowledge and spatial reasoning skills that LLMs acquire from their extensive training data as a possible means of developing a geo-foundation model [7][8][9]. A wide variety of spatial data sets are readily accessible from a number of different sources and sizes, such as trajectory data, aerial images, geotags created by the public, results from motion sensors, and film from dashboard cameras [10][11]. Despite having undoubtedly encountered geographical information during training, the extent to which general-purpose LLMs can reason about implicit spatial correlations is unknown [12]. The importance of this subject is growing as LLMs are used for more and more complicated tasks, some of which have a physical basis, such as creating paths between known sites or recommending destinations to users depending on their position and trajectory.

This paper assesses the geospatial reasoning ability of LLMs through experiments covering a broad range of spatial tasks, including toponym resolution and reasoning about four fundamental spatial relations: metric, directional, topological, and order relationships. Previous work has shown that LLMs possess basic spatial awareness [13][14], such as knowledge of geocoordinates, directional relationships between major cities, and distances between cities [15][16]. However, by extending the tasks to cover all major spatial relations and increasing task complexity to involve multiple spatial entities, LLMs perform poorly, especially in complex spatial reasoning[17]. This study is motivated by the need to rigorously assess and improve LLMs' ability to reason about space through unbiased, controlled experiments, ultimately pushing the boundaries of their capabilities in real-world, geospatial tasks.

- A new evaluation dataset centered on Australian geography was created to minimize the influence of pre-training memorization in LLMs, offering a more rigorous and unbiased assessment of spatial reasoning capabilities.
- The study systematically assessed LLMs across five key spatial reasoning tasks Toponym Resolution, Metric Relations, Directional Relations, Topological Relations, and Cyclic Order Relations providing a detailed, task-specific understanding of model strengths and weaknesses.
- By maintaining zero-shot prompting, fixed temperatures, and constant random seeds, the research ensured highly controlled experimental conditions, enabling reproducibility and reducing evaluation noise commonly seen in LLM benchmarking.
- The observation that certain models (e.g., from Google and Anthropic) displayed higher abstention rates highlights important differences in model confidence

calibration and strategies for handling spatial uncertainty in complex reasoning tasks.
- The study uncovered major deficiencies in LLMs' ability to perform multi-entity and sequential spatial reasoning, especially in cyclic order tasks and nuanced metric relation interpretation, providing critical insights for future model improvement and spatial reasoning research.

This is how the remainder of the paper is structured. Recent study evaluating LLMs' geographical knowledge is described in Section II. The required grounding in spatial thinking is provided in Section III. The experimental procedure is described in Section IV, the findings are shown in Section V, and the commentary is presented in Section VI. Section VII concludes with a discussion of future work in Section VIII.

II. Related Work

Many recent works have explored the abilities and limitations of LLMs for performing various reasoning tasks, from math word problems [3] to place name resolution [18]. This section summarizes recent work investigating the extent to which LLMs can reason spatially. It begins by describing studies probing LLMs for spatial and geographic knowledge, followed by a survey of recent efforts evaluating and enhancing spatial reasoning in large language models. papers on geo-foundation models and finally discuss specific techniques proposed to enable LLMs to ingest and reason over spatial data.

A. General Spatial and Geographic Knowledge In LLMS

Recent studies have examined the capabilities of Local Machine Learning (LLM) tools like ChatGPT in answering spatial-related questions. They found limitations in ChatGPT's GIS knowledge [19], unreliability in general reasoning tasks, and failures in generic spatial reasoning questions. LLMs also performed poorly in planning tasks involving spatial reasoning, such as moving objects for robotic applications. Additionally, faults were found in map visualizations and sketch maps generated by ChatGPT using code or ASCII symbols.

B. Geospatial LLMS and Geo-Foundation Models

Two vision papers [20] suggest that Local Machine Learning (LLM) has potential for geospatial databases, performing spatial reasoning with natural language prompts. Experiments show accurate geocoordinates and names of cities near or far from a reference city [21]. However, LLM cannot adapt to scales. The authors propose a geo foundation model pre-trained on different data modalities and a generic foundation model for human mobility data at various scales [22]. Future work will address challenges in embeddings, model architectures, and self-supervised tasks.

C. LLM adaptations for Geodata

A few embedding methods and model architectures have been proposed to enable LLMs to handle certain types of geospatial data. For trajectories of geocoordinates, [23] propose an embedding method that uses sub-trajectory similarity learning to pre-train trajectory representations that can be used in downstream prediction tasks. For textual georeferences, [24] design a pre-training task using spatial coordinate embeddings (based on latitude and longitude) corresponding to textual georeferences, which improve accuracy over non-spatial methods on the downstream tasks of geoentity type prediction and linking to knowledge graphs. For spatio-temporal forecasting [25], proposes a method of tokenization and encoding to increase LLM understanding of spatio-temporal text references. While significant headway has been made recently in embedding geodata, there remains a lack of self-supervised training objectives pertaining to geospatial reasoning, which discuss further in section VII.

III. Spatial Reasoning

This paper defines spatial reasoning, a task involving understanding spatial relationships between entities or objects in space. It discusses the challenges of addressing spatial reasoning using LLMs, which are divided into mereology, topology, and location proper theories. Spatial reasoning tasks involve answering questions about spatial relations between objects in space.

A. Spatial Entities

Spatial entities are fundamental elements of spatial data, categorized into three types: points (x, y) in Cartesian space, lines (shortest path among points) [26], and regions (polyline joining points). Points represent locations, lines represent ways, and regions represent areas. Spatial entities are essential in understanding the physical world and are often used in LLM questions[27].

B. Spatial Relations

The following kinds of connections describe the spatial relationships between points, lines, and regions [25]:

- Distances between geographical objects are described by metric relations, which may be either quantitative (like "ten miles") or qualitative (like "near" or "far").
- Topological relations that describe how regions, lines, and points interact (‘equals’, ‘disjoint,’ ‘intersects,’ ‘touches,’ ‘partially overlaps’, ‘within,’ ‘contains’) [28],
- Directional relations (such as "North," "Left," or "Behind") that characterize an entity's relative location in space and
- Order relations that describe the cyclic order in which objects appear with respect to a central coordinate[29] (‘clockwise’ or ‘counterclockwise’).

Spatial relations enable qualitative or quantitative descriptions of how physical places or objects in the world interact spatially. The ability to correctly identify spatial elements and understand the relationships between them is essential for LLMs to provide proper answers to spatial queries [30]. For example, city A is North of city B, and city B is North of city C; therefore, city A is north of city C.

IV. Methodology

This section describes the procedure for testing LLMs' spatial thinking skills, including the models and questions that were used. A series of experiments were designed to assess spatial reasoning through toponym resolution and four fundamental spatial relations: metric, directional, topological, and order relationships.

A. Dataset

The evaluation dataset was created to minimize the impact of specific questions on pre-training models. Australia was chosen due to its English-speaking population and less widely documented place names [31]. To minimize toponym resolution, comma groups were used in questions. The dense

population and wide-open spaces of Australia made it possible to put the models through their paces using a mix of long and short distance testing. Dual-naming locations allowed for tokens less likely to be memorized in geospatial contexts. The dataset contains 239 questions covering point, line, region, metric, directional, topological, and cyclic relations commonly used in spatial pattern matching.

TABLE I. SUMMARY OF MODELS EVALUATED.

Developer	Model	#Param $	per M/Tok
OpenAI	gpt-3.5-turbo	†	00.50/01.50
	gpt-4	†	30.00/60.00
	gpt-4-turbo	†	10.00/30.00
	gpt-4o	†	05.00/15.00
Google	gemini-1.0pro	†	00.50/01.50
	gemini-1.5-flash	†	00.35/01.05
	gemini-1.5-pro	†	03.50/10.50
Anthropic	claude-3-opus	†	15.00/75.00
	claude-3-sonnet	†	03.00/15.00
	claude-3-haiku	†	00.25/01.25
Meta	llama3-70b	70b	03.20/03.20
	llama3-8b	8b	01.60/01.60
Mistral	mixtral-8x22b-instruct	39b/141b	03.20/03.20
	mistral-7b-instruct	7b	01.60/01.60

B. Models

Fifteen models were selected from five leading developers of Large Language Models, covering a range of parameter sizes or self-reported capabilities when parameter data was unavailable. These models were accessed through their Application Programming Interfaces. Table I summarizes the chosen models together with their parameters and the estimated cost of usage. To conduct experiments, it reduced the temperature of each model to zero and used consistent seed values wherever possible to reduce the effect of generational randomness.

C. Prompting

The model being tested is given each question as a separate prompt, guaranteeing that no context is carried over from one encounter to another. Zero-shot prompting approaches are used, except when shaping the output format or in metric experiments that aim to elicit reasoning through in-context learning. The purpose of each exercise is to assess spatial thinking, and the first prompt for each is as follows:

You are answering to evaluate spatial reasoning ability. You will be presented a question and asked to answer. Where there are multiple possible answers, select the most likely. Answer as briefly as possible, preferring single-word answers where they suffice. Where do not know the answer, it is unanswerable or you are uncertain, return' ICATQ'.

Prompt 1: Initial System Prompt

A current experiment dictates the following prompt. I fill in the exact letters (A, B, C, etc.) that represent geographical elements (such as cities, rivers, highways, and states) to create the final prompts for each experiment. This structure is described below for the second prompt.

Experiment 1: Toponym Resolution

Where is A? Format your answer as a comma-separated list: state/county, country.

Prompt 2: Toponym Resolution Prompt

The dataset is tested using unaided toponym resolution to identify the strong association of each term with Australia. Points are awarded for linking it to Australia and additional points for complete comma groups. To provide room in the dataset for contextual spatial reasoning, the toponym experiment removes phrases closely linked to other nations, leaving only unresolved toponyms. Subsequent trials use comma groups for site specification in an effort to mitigate the effect of toponym resolution on the assessment of spatial reasoning [32][33].

Experiment 2: Metric Relations

The purpose of this research is to develop three different kinds of prompts and see if LLM models can reason about metric spatial links. A neutral city whose distance from C is comparable to that of A and B is requested in the first neutral prompt.

The distance from A to B is similar to the distance from C to what other city or town?

Prompt 3: Neutral Metric Prompt

Based on previous studies [20], LLMs are shown to find places closer to a query location when the prompt includes the word "near," and destinations farther distant from the query location when the prompt includes the word "far." There are two versions of the neutral metric prompt that have been developed to investigate this topic further. Incorporating the terms "near" and "far" into the question is an attempt to gauge whether LLMs retain a fixed understanding of those terms or whether it can modify their meaning according to the question's scale. The 'near' and 'far' prompts are as follows:

If A is ' near' to B, what is a city or town 'near' to C?

Prompt 4: 'Near' Metric Prompt

If A is 'far' from B, what is a city or town 'far' from C?

Prompt 5: 'Far' Metric Prompt

Place names from all throughout Australia, including both Indigenous and Western ones, should be used to populate the prompts. As a measure of expected distance, find the geodesic distance between the LLM's returned place and the query placement C. Ideally, the distance x, which is the stated goal distance, and the projected distance will be quite nearby. Both distances are normalized by the country's approximate diameter, so error tolerance scales with x.

D. Experiment 3: Directional Relations

To determine if LLMs can reason about the spatial relationships between multiple locations, construct a series of prompts asking about the cardinal directionality between pairs of entities and create 42 queries covering 18 Australian city and town names of varying population $size^1$ into groups of '2-way' and '3-way' constraint problems. For each group consisting of entities A, B, and C, construct the following 2-way and 3-way directional prompts:

A is north, northeast , northwest , south , southeast , southwest , east , or west of B?

Prompt 6: 2-way Directional Prompt

A is north , northeast , northwest , south , southeast , southwest , east , or west of B and C?

Prompt 7: 3-way Directional Prompt

Repeat this for each permutation of A, B, and C, reordering them within the prompt text. Scoring for directional

relations rewards specificity, with more points being awarded for ‘northwest’ rather than ‘north’ or ‘west’ when both could be true.

E. Experiment 4: Topological Relations

To determine if LLMs can reason about topological relations, constructed a series of seven topological relation prompts containing questions pertaining to each of the major topological predicates and select points (P) from a set of city and town names in Australia, regions (R) from a set of lakes, parks, regions, and states in Australia, and lines (L) from a set of highways, roadways, and riverways in Australia.

The prompts are structured as follows:

Ranging from 5,297,089 to 37

1. Is A geospatially equal to B ?
2. Is A geospatially disjoint from B ?
3. Does A geospatially intersect B ?
4. Does A geospatially touch B ?
5. Does A geospatially partially overlap B ?
6. Is A geospatially within B ?
7. Does A geospatially contain B ?

Prompts 8-14: Topological Relation Prompts

populates the prompts with cities and towns of varying populations, major waterways, and the ‘common name’ for major roadways (i.e.,”The Pacific Highway” rather than ”M1 Motorway”). The sampling was conducted across the states and territories of Australia, including both indigenous and western place names. The binary response 'Yes' or 'No' is used to score each answer.

F. Experiment 5: Cyclic Order Relations

To determine if LLMs can reason about cyclic order relationships, construct a series of prompts asking about the clockwise or counterclockwise relationship between entities. For each group consisting of entities A, B, and C, and construct the following prompt:

With respect to a centroid in A, is moving from B to C a clockwise or counterclockwise direction?

Prompt 15: Cyclic Order Relation Prompt

Permute the ordering of A, B, and C within the prompt text and measure the binary response ‘clockwise’ or ‘counterclockwise’ for each answer.

V. Results

The results reveal variations in model performance across different spatial reasoning tasks. Metric relation tests showed significant errors in predicted distances, particularly with the “near” keyword. Directional relations indicated that GPT models performed better, but accuracy dropped with increased constraints. Topological queries outperformed other types, though line-based relations had higher errors. Cyclic order prompts performed at random levels, highlighting reasoning limitations. Across all tasks, Google and Anthropic models exhibited higher abstention rates, reflecting uncertainty in complex spatial queries. To discuss the outcomes of the experiments described in the previous section.

A. Result 1. Metric Relations

Different kinds of prompts had different distances between their expected and real locations, as shown in Experiment 2. There were several instances when the distances were off by hundreds of km. The algorithms repeatedly selected locations that were too close to a query point, particularly when employing a term "near," leading to locations that were never more than 1,000 km far.

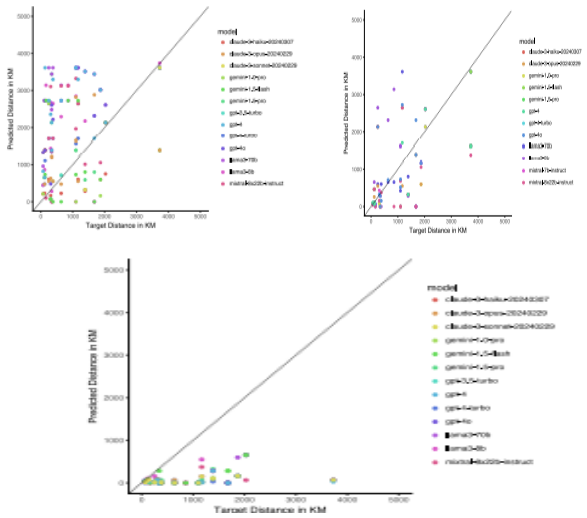

Fig. 1. Target vs. Predicted Distance for Far, Neutral Metrics and near’

a) The 'near' metric prompt's estimated distance compared to its target distance. Out of the 280 replies from the model, 89 did not participate, and eleven were outliers with very high projected distances.
b) The 'far' metric prompt's estimated distance compared to its target distance. Out of 280 answers from the model, 100 were abstentions and three were outliers with very high anticipated distances.
c) The neutral metric prompt's target distance compared to its expected distance. It excluded four extreme outliers with very large estimated distances and 159 non-respondents from the total of 280 model runs.

Figure 1 displays the outcomes of the 'far,' 'neutral,' and 'near' metric connection prompts. 'A' and 'B' make up the goal distance in Prompts 3-5, whereas 'C' and the location given by the model make up the projected distance. When using metric spatial reasoning, it is more accurate to place points closer to the line drawn at y= x than further away.

B. Result 2. Directional Relations

The results for the 22 2-way directional prompts in Figure 2(a) varied across the models, with Claude-Haiku unable to answer any question and mistral-7b again struggling because of token over-generation. The tests for the 20 3-way prompts summarized in Figure 2(b) show that half of the models tested showed a significant increase in model abstention and error rate. For comparison [13], performed pairwise directional prompting for major cities in Australia and found the responses to be correct in 44 out of 50 cases.

C. Result 3. Topological Relations

Topological queries generally outperform other relation types, but line-based relations have higher error rates due to reduced terms in training. The better performance is attributed to qualitative nature and limited test involvement [34].

D. Result 4. Order Relations

The returned results for the cyclic order relation prompts (Figure 4) show performance on par with random guessing between ‘clockwise’ and ‘counterclockwise’ by the models.

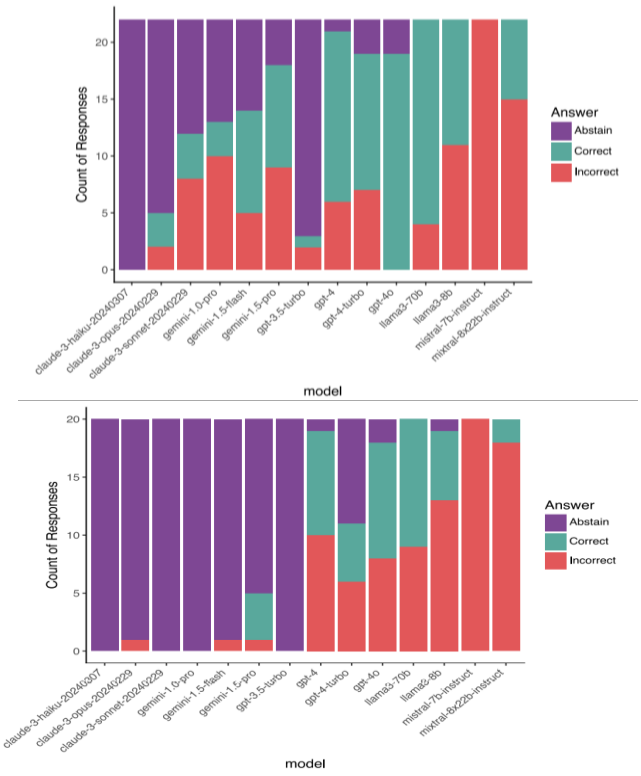

Fig. 2. Two directional relation prompts 2-way and 3-way

a) Model performance on 2-way directional relation prompts. 2-way directional relations show gpt family of models are stronger directional reasoners than other models.
b) Model performance on 3-way directional relation prompts. A third constraint decreases answer confidence and accuracy.

Figure 2 shows the results of two directional relation prompts 2-way and 3-way. Adding a third directional constraint reduces model performance, and also decreases the chance that the question was observed in an LLM's training data.

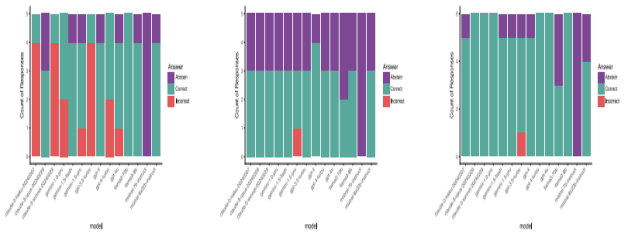

Fig. 3. Error Patterns and Reasoning Challenges in Geospatial Queries

a) Higher error rates occur in line-based queries across the evaluation set, compared to points and regions.
b) Partial Overlap relations are primarily border regions and introduce uncertainty about ownership, which is reflected in the higher-than-average abstention rate across all models.
c) The GPT-3.5-turbo's error came from a question about whether Canberra is within the state of NSW. The physical sense contradicts the political in this instance, highlighting some of the reasoning difficulties faced in geospatial computing.

Figure 3 shows a selection of topological results reflects broader trends in their evaluation of line-based queries performing worse than region or point-based and a preference towards abstaining from answering under uncertainty.

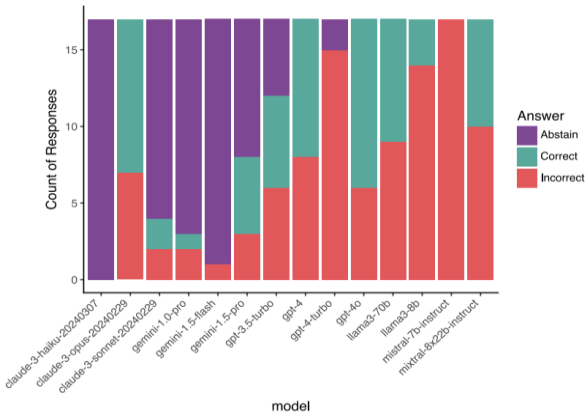

Fig. 4. Results of cyclic order relation prompt

Figure 4 displays the outcomes of cyclic order relation prompt. The results for cyclic order relation prompts indicate that model performance is comparable to random guessing between 'clockwise' and 'counterclockwise.' This suggests that the models struggle with accurately determining cyclic order relationships, highlighting a limitation in their reasoning capabilities for such tasks. Across the experiments, the Google and Anthropic models consistently abstained at higher rates.

VI. Discussion

The study highlights key insights into LLMs' spatial reasoning abilities across different relational metrics. Findings indicate that LLMs associate the word "near" with smaller distance metrics and "far" with larger ones, suggesting a static interpretation of proximity. Directional reasoning shows a gap between pairwise and three-way relations, revealing limited spatial inference skills beyond memorized data. Topological relations, particularly for line entities, show poor performance, likely due to the lower prevalence of such relationships in training data compared to widely known cities and regions. Similarly, order relations exhibit weak model performance, with LLMs struggling to reason about relative city positions, indicating limited exposure to geospatial reasoning.

VII. Future Work

This section discusses ways to improve the spatial reasoning abilities of LLMs by explicitly devising embedding techniques and self-supervised training objectives that align with spatial tasks.

A. Embedding Techniques

The first envisages the development of novel embedding methods for various types of geodata. To allow LLMs to learn complex spatial relations, such as multi-way directional relations that it shown to be a shortcoming in Experiment 3, it proposes using an appropriate encoding scheme for that type of information. In the spatial pattern matching domain, complex spatial relationships are captured using graph encodings, where relations can be made explicit using the edges between graph nodes [22]. With the data in this format, spatial reasoning can be formulated as graph reasoning, which can be captured in a learning objective [35].

B. Self-Supervised Training Tasks

Self-supervised training objectives are crucial for learning embedding methods during pre-training. Initial methods for pre-training spatial coordinate embeddings improve entity type classification and linking tasks. However, more intuitive tasks capturing two-dimensional spatial relationships are needed. Leveraging natural language descriptions and logical reasoning is needed for complex spatial questions. Self-supervised training objectives can be generated through programmatic generation of fictitious worlds [24].

C. Long Term Opportunities: Multimodal Spatial Learning

Once LLMs can be successfully pre-trained on various types of spatial data with success in downstream tasks, it will then be possible to leverage work in multimodal learning to combine a variety of input modalities of geospatial data. Doing so would enable the training of a more generic geo-foundation model broadly capable of spatial reasoning given new sources of spatial information, at a variety of scales. To accomplish this goal, multimodal learning techniques that have been successful at enabling question answering over visual, textual, and other types of data could provide benefits in the spatial domain [21].

VIII. CONCLUSION

Recent work has demonstrated that large language models (LLMs) have some level of spatial awareness, such as knowledge about geocoordinates, directional relationships between major cities, and relative distances between cities. This study highlights the varying performance of large language models (LLMs) across different spatial reasoning tasks. The models demonstrated relative strength in tasks involving qualitative spatial relations, such as topological reasoning, while facing challenges in more quantitative tasks like metric distance estimation and cyclic order reasoning. The results underscore the potential of LLMs in certain spatial contexts, with topological relations being a particular area of strength, while other tasks revealed areas for improvement. Overall, the evaluation provides valuable insights into the spatial reasoning capabilities of LLMs and contributes to the understanding of their effectiveness in handling complex geospatial queries. This study has several limitations, including its focus on a specific geographic region (Australia) and a limited set of predefined spatial reasoning tasks, which may not fully capture the diverse challenges LLMs face in real-world spatial reasoning. Additionally, the use of zero-shot prompting may not exploit the models' full potential, particularly when fine-tuned for specific tasks. Future work could address these limitations by expanding the dataset to cover a broader range of geographic contexts and spatial tasks, incorporating fine-tuned models, and integrating external geospatial data for improved accuracy. Exploring advanced reasoning strategies, such as multi-step reasoning and better uncertainty handling, could also enhance LLMs' performance in complex spatial scenarios.

REFERENCES

[1] R. Kumar, “Leveraging LLMs for Continuous Data Streams_ Methods and Applications,” *ICIDA*, 2025.

[2] S. Pahune and M. Chandrasekharan, “Several categories of Large Language Models (LLMs): A Short Survey,” *Comput. Sci. > Comput. Lang.*, Jul. 2023, doi: 10.22214/ijraset.2023.54677.

[3] G. Badaro, M. Saeed, and P. Papotti, “Transformers for Tabular Data Representation: A Survey of Models and Applications,” *Trans. Assoc. Comput. Linguist.*, vol. 11, pp. 227–249, 2023, doi: 10.1162/tacl_a_00544.

[4] M. Duckham *et al.*, “Qualitative spatial reasoning with uncertain evidence using Markov logic networks,” *Int. J. Geogr. Inf. Syst.*, vol. 0, pp. 1–34, 2023, doi: 10.1080/13658816.2023.2231044.

[5] N. R. Saurabh Pahune, “Large Language Models and Generative AI’s Expanding Role in Healthcare,” *Researcg Gate*, 2024.

[6] S. Pandya, “Comparative Analysis of Large Language Models and Traditional Methods for Sentiment Analysis of Tweets Dataset,” *Int. J. Innov. Sci. Res. Technol.*, vol. 9, no. 12, pp. 1647–1657, 2024, doi: https://doi.org/10.5281/zenodo.14575886.

[7] G. Mai *et al.*, “On the Opportunities and Challenges of Foundation Models for Geospatial Artificial Intelligence,” vol. 1, no. 1, pp. 1–37, 2023.

[8] A. Immadisetty and J. Olusegun, “Real-Time Data Analytics in Customer Experience Management: A Framework for Digital Transformation and Business Intelligence,” *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 10, no. 6, pp. 1280–1288, 2021.

[9] H. Xue and F. Salim, “Artificial General Intelligence for Human Mobility (Vision Paper),” 2023, pp. 1–4. doi: 10.1145/3589132.3625652.

[10] R. Kumar, “Evaluating and Enhancing Spatial Reasoning in Large Language Models,” *ICIDA*, 2025.

[11] S. G. Prity Choudhary, Rahul Choudhary, “Enhancing Training by Incorporating ChatGPT in Learning Modules: An Exploration of Benefits, Challenges, and Best Practices,” *Int. J. Innov. Sci. Res. Technol.*, vol. 9, no. 11, 2024.

[12] S. Murri, “Graph Database Pruning for Knowledge Representation in LLMs,” *dzone*, 2025, [Online]. Available: https://dzone.com/articles/graph-database-pruning-for-knowledge-representation-in-llms

[13] J. Qi, Z. Li, and E. Tanin, “MaaSDB: Spatial Databases in the Era of Large Language Models (Vision Paper),” *GIS Proc. ACM Int. Symp. Adv. Geogr. Inf. Syst.*, no. 2, 2023, doi: 10.1145/3589132.3625597.

[14] J. H. Lee, M. Sioutis, K. Ahrens, M. Alirezaie, M. Kerzel, and S. Wermter, “Neuro-symbolic spatio-Temporal reasoning,” *Front. Artif. Intell. Appl.*, vol. 369, pp. 410–429, 2023, doi: 10.3233/FAIA230151.

[15] S. Pahune, Z. Akhtar, V. Mandapati, and K. Siddique, “The Importance of AI Data Governance in Large Language Models,” *Preprints*, Apr. 2025, doi: 10.20944/preprints202504.0219.v1.

[16] S. Nokhwal, P. Chilakalapudi, P. Donekal, S. Nokhwal, S. Pahune, and A. Chaudhary, “Accelerating Neural Network Training: A Brief Review,” *ACM Int. Conf. Proceeding Ser.*, pp. 31–35, 2024, doi: 10.1145/3665065.3665071.

[17] S. Pandya, “Comparative Analysis of Large Language Models and Traditional Methods for Sentiment Analysis of Tweets dataset for text classification,” *Int. J. Innov. Sci. Res. Technol.*, vol. 9, no. 12, 2024.

[18] C. V. A. Minervino, C. E. C. Campelo, M. G. de Oliveira, and S. D. Silva, “QQESPM: A Quantitative and Qualitative Spatial Pattern Matching Algorithm,” *Proc. Brazilian Symp. GeoInformatics*, pp. 49–60, 2023.

[19] Y. Bang *et al.*, “A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity,” pp. 675–718, 2024, doi: 10.18653/v1/2023.ijcnlp-main.45.

[20] P. Bhandari, A. Anastasopoulos, and D. Pfoser, “Are Large Language Models Geospatially Knowledgeable?,” in *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems*, in SIGSPATIAL ’23. New York, NY, USA: Association for Computing Machinery, 2023. doi: 10.1145/3589132.3625625.

[21] N. Fei *et al.*, “Towards artificial general intelligence via a multimodal foundation model,” *Nat. Commun.*, vol. 13, no. 1, pp. 1–13, 2022, doi: 10.1038/s41467-022-30761-2.

[22] C. Strobl, “Dimensionally Extended Nine-Intersection Model (DE-9IM),” in *Encyclopedia of GIS*, S. Shekhar and H. Xiong, Eds., Boston, MA: Springer US, 2008, pp. 240–245. doi: 10.1007/978-0-387-35973-1_298.

[23] M. Hu *et al.*, “Self-supervised Pre-training for Robust and Generic Spatial-Temporal Representations,” *Proc. - IEEE Int. Conf. Data Mining, ICDM*, pp. 150–159, 2023, doi: 10.1109/ICDM58522.2023.00024.

[24] Z. Li, J. Kim, Y.-Y. Chiang, and M. Chen, "{S}pa{BERT}: A Pretrained Language Model from Geographic Data for Geo-Entity Representation," in *Findings of the Association for Computational Linguistics: EMNLP 2022*, Y. Goldberg, Z. Kozareva, and Y. Zhang, Eds., Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, Dec. 2022, pp. 2757–2769. doi: 10.18653/v1/2022.findings-emnlp.200.

[25] L. Liu, S. Yu, R. Wang, Z. Ma, and Y. Shen, "How Can Large Language Models Understand Spatial-Temporal Data?," vol. 1, 2024.

[26] P. Bertella, Y. Lopes, R. Oliveira, and A. Chaves Carniel, "A Systematic Review of Spatial Approximations in Spatial Database Systems," vol. 13, pp. 224–240, 2022, doi: 10.5753/jidm.2022.2519.

[27] RAJARSHI TARAFDAR, "Algorithms on Majority Problem," *Univ. Missouri-Kansas City*, p. 6, 2015.

[28] E. Clementini, J. Sharma, and M. J. Egenhofer, "Modelling topological spatial relations: Strategies for query processing," *Comput. Graph.*, vol. 18, no. 6, pp. 815–822, 1994, doi: https://doi.org/10.1016/0097-8493(94)90007-8.

[29] A. Schwering, J. Wang, M. Chipofya, S. Jan, R. Li, and K. Broelemann, "SketchMapia: Qualitative Representations for the Alignment of Sketch and Metric Maps," *Spat. Cogn. Comput.*, 2014, doi: 10.1080/13875868.2014.917378.

[30] H. Chen, Y. Fang, Y. Zhang, W. Zhang, and L. Wang, "ESPM: Efficient Spatial Pattern Matching (Extended Abstract)," 2020, pp. 2038–2039. doi: 10.1109/ICDE48307.2020.00238.

[31] A. C. Carniel, "Spatial Information Retrieval in Digital Ecosystems: A Comprehensive Survey," in *Proceedings of the 12th International Conference on Management of Digital EcoSystems*, in MEDES '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 10–17. doi: 10.1145/3415958.3433038.

[32] Saransh Arora and Sunil Raj Thota, "Using Artificial Intelligence with Big Data Analytics for Targeted Marketing Campaigns," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 593–602, Jun. 2024, doi: 10.48175/IJARSCT-18967.

[33] M. D. Lieberman, H. Samet, and J. Sankaranayananan, "Geotagging: using proximity, sibling, and prominence clues to understand comma groups," in *Proceedings of the 6th Workshop on Geographic Information Retrieval*, in GIR '10. New York, NY, USA: Association for Computing Machinery, 2010. doi: 10.1145/1722080.1722088.

[34] A. Chaves Carniel, "Defining and designing spatial queries: the role of spatial relationships," *Geo-spatial Inf. Sci.*, vol. 27, pp. 1–25, 2023, doi: 10.1080/10095020.2022.2163924.

[35] N. Schneider, K. O'Sullivan, and H. Samet, "The Future of Graph-based Spatial Pattern Matching (Vision Paper)," 2024, pp. 360–364. doi: 10.1109/ICDEW61823.2024.00054.

