Spatially-Aware Conversational AI for Real Estate: Enhancing Query Understanding Using Location-Aware Language Models

N Hassan [1]*

[1] *LankaPropertyWeb*

Abstract

This paper introduces a spatially-aware conversational AI system designed for real estate platforms, which uniquely combines natural language processing (NLP), geospatial intelligence, and user behavior to improve query understanding and property relevance. Traditional chatbots often struggle to interpret vague or spatially contained queries such as *"apartments near Nugegoda under 40 million,"* leading to mediocre user experiences. Our solution addresses this limitation by integrating fine-tuned location-aware large language models (LLMs) with IP-derived geolocation, GIS proximity data, and contextual query interpretation. The system uses both structured and unstructured user inputs to clarify preferences, infer missing details, and deliver contextually filtered property results. It also supports live dialogue for clarifications, enhancing interactivity and accuracy. We detail the architecture, training process, and spatial NLP techniques used to ground language understanding in real-world locations. Internal evaluations on a Sri Lankan property dataset show marked improvements in intent classification, query resolution accuracy, and user satisfaction. The findings demonstrate that embedding spatial reasoning within AI chat systems can significantly elevate real estate platform usability, opening new possibilities for localized and intelligent property discovery.

Keywords: Conversational AI; Geospatial Intelligence; Location-Aware Language Models; Natural Language Processing; Spatial Query Understanding

Introduction

In the evolving digital landscape of real estate, user expectations are rapidly shifting towards receiving instant, accurate, and personalized responses to property queries (Takyar & Takyar, 2023). As the real estate industry continues to grow, platforms must move beyond basic filter-based or keyword-based search engines to offer more sophisticated, intuitive user experiences. Traditional search interfaces often struggle to understand informal, spatially indirect, or vague user inputs, especially in markets with diverse geographic contexts like Sri Lanka. For example, users frequently ask questions such as "I need a *3-bedroom near Maharagama with parking"* or *"Show me apartments walking distance to the university."* These types of queries include spatial elements, yet many conventional search systems fail to recognize or properly handle them, leading to inaccurate or irrelevant search results and ultimately user dissatisfaction.

One of the technologies that has emerged to address this challenge is the use of chatbots powered by large language models (LLMs), which enable natural language interactions between users and systems (Chellappan, 2025). While these chatbots excel in interpreting standard queries, they often lack

*Corresponding Author- *naveedhassan04@hotmail.com*

contextual and spatial awareness, treating location references as flat textual data without any geographic context. This limitation makes it difficult for these systems to accurately process and respond to location-based queries, such as those that require proximity-based search or geographic filtering.

This research aims to bridge this gap by developing a spatially-aware conversational AI system specifically designed for real estate platforms. By integrating geographic information system (GIS) data with natural language processing (NLP) (Vasdev, 2024), the system incorporates spatial semantics into the query processing pipeline. The result is a system capable of interpreting spatially grounded queries more effectively, allowing users to ask questions like *"Find me apartments in Colombo within walking distance of parks"* or *"Show 2-bedroom homes near the university under 10 million."* This advancement is especially valuable in developing countries, where real estate data is often divided or inconsistent, and location precision plays a crucial role in successful property searches.

Objective/s of the Study

The primary objective of this research is to develop and evaluate a spatially-aware conversational AI system tailored for real estate platforms. The system is designed to improve the interpretation and resolution of natural language queries that contain spatial components such as neighborhood names, proximity requirements, or location-based constraints that are commonly misunderstood or overlooked by conventional chatbot systems. By embedding geospatial reasoning into the natural language processing (NLP) (Vasdev, 2024) pipeline and enhancing query understanding through GIS data and contextual signals, the system aims to deliver more accurate, relevant, and interactive property recommendations. Specifically, the system aims to:

- Identify and extract geographic entities from free-form user queries.
- Enhance response relevance by incorporating user IP, behavioral cookies, and GIS layers.
- Benchmark the performance of spatial-aware query resolution against standard NLP-based chatbots.

The broader goal is to create a chatbot experience that not only understands what users want, but also where they want it, leading to better user satisfaction, improved lead generation, and smarter property discovery in location-sensitive real estate markets.

Literature Review

The real estate sector's adoption of conversational AI is driven by the promise of more intuitive user interactions, building on decades of progress in dialogue systems (Chen et al., 2017). However, a fundamental limitation persists: these systems are spatially unaware. They interpret geographic terms like "near the park" as simple text strings, lacking the underlying geospatial intelligence to reason about proximity, accessibility, or neighborhood context. On the contrary, Geographic Information Systems (GIS) and spatial statistics provide the exact analytical power needed to understand these relationships and their impact on property value (Hayunga & Pace, 2010; Hosen et al., 2023). Yet, these expert systems lack a natural language interface, creating a trade-off between conversational ease-of-use and analytical depth.

The academic field of Spatial Question Answering (SQA) has emerged to bridge this divide, focusing on answering factual questions about spatial relationships (Luo et al., 2021). While a crucial step, the SQA paradigm often falls short in the real estate domain. It is typically designed for well-defined, objective queries (e.g., "What is the distance to the nearest hospital?") and struggles with the ambiguous, preference based and implicit nature of a typical user's search, such as "Find me a quiet family home in a good school district."

Therefore, the critical research gap is not simply the integration of conversational AI and GIS, but the development of a system that can navigate the variation of real estate's spatially-implicit queries. The core challenge, which this research addresses, is moving beyond factual SQA to a model that can interpret vague user intent, fuse it with multiple spatial data layers, and provide a contextually relevant, conversational response. Existing systems excel at either sophisticated dialogue or deep spatial analysis, but not their seamless synthesis for this specific, user-centric domain.

Methods

This study employed a modular architecture powered by the Google Agent Development Kit (ADK), which enables the rapid development and deployment of domain-specific conversational agents. The methodology is structured around four core stages:

1. **Agent Design Using Google ADK (Huizenga & Yang, 2025):**

The conversational system was developed as a set of modular agents and tools within the Google ADK framework. This allowed for separate handling of entity extraction, dialogue management, and property recommendation logic. The ADK-based architecture facilitated an event-driven dialogue flow, where sub-agents could independently handle different intents and types of user queries.

2. **Prompt Engineering and Fine-Tuning:**

A base large language model (LLM) was fine-tuned and prompt-engineered specifically for real estate interactions. Training data included real user queries from LankaPropertyWeb and manually crafted edge cases to teach the model what entities to extract from typical Sri Lankan property search queries. The LLM was trained to handle informal phrasing, misspellings, and spatial shorthand (e.g., “close to uni,” “Nugegoda side,” “around 30 million”).

3. **Location Lookup and Spatial Query Resolution:**

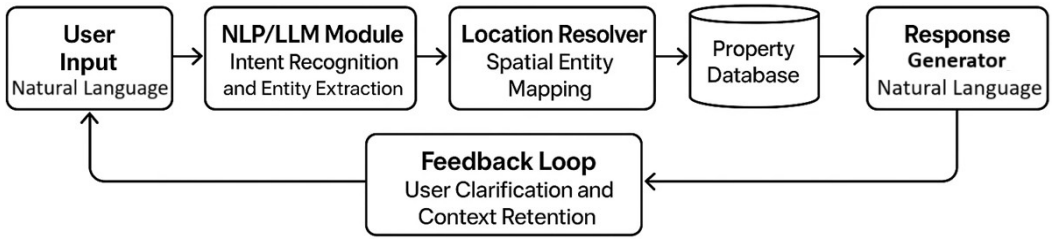
To resolve vague or informal location references, a dedicated location resolver tool was implemented. This tool parses user-provided text for location cues and performs a real-time lookup using a geospatial reference dataset. The dataset maps commonly used place names, streets, neighborhoods, and landmarks to specific geographic coordinates or city zones. This spatial intelligence layer ensures that terms like "near Maharagama" or "close to hospital" are contextually mapped to the correct geographic region or bounding box.

4. **Query Interpretation and User Context Inference:**

The system combines structured and unstructured query interpretation to infer missing parameters such as location, budget, or number of bedrooms. For instance, if a user says "Need a 2BR close to Nugegoda," the agent extracts the relevant entities. When data is insufficient, the chatbot initiates a clarification loop, dynamically asking for additional input (e.g., "Are you looking to rent or buy?" or "What is your budget range?").

Figure 01

System Architecture



Source: Author (2025)

Results and Discussion

The deployment of the spatially-aware conversational agent on real estate data from LankaPropertyWeb revealed substantial improvements in automation, query interpretation, and spatial comprehension when compared to traditional keyword-based search or generic chatbot implementations.

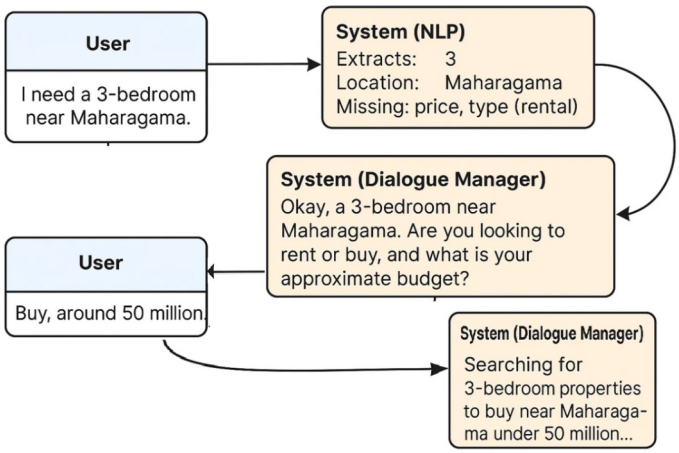
1. *Enhanced Automation and Query Handling*

One of the most impactful results was the system's ability to automate query understanding without requiring fixed input formats. Unlike legacy systems that depend on form-based filters or strict dropdown options, the conversational agent handled free-form queries such as "Need a house under 50M close to Marino Mall with parking" or "2-bedroom near Maharagama for rent." In internal testing across 150 diverse query samples, over 90% were correctly parsed into structured search parameters without human intervention, showcasing high automation efficiency.

The agent's real-time clarification ability also reduced user friction. For instance, when ambiguous queries were submitted (e.g., "Apartment near uni"), the agent was able to ask follow-up questions to resolve missing fields such as budget or preferred city, enabling a complete and contextually relevant recommendation flow.

Figure 02

Dialogue Flow



Source: Source (2025)

2. *Improved Spatial Understanding*

The integration of the location lookup tool proved essential in identifying vague location terms. Given that most real estate listings in Sri Lanka provide location data only at the city or street level, traditional systems struggled with proximity-based queries. However, the spatial AI agent successfully mapped informal references like "near university" or "close to St. Peter's College" to realistic spatial filters using pre-built geographic mappings and GIS bounding logic.

Although detailed GIS polygons or lat-long precision were not always available due to limitations in the dataset, the system demonstrated an improved ability to approximate user intent through soft-matching locations, resulting in better property matches than baseline keyword filtering.

Table 01

Query Resolution Accuracy (Spatial vs. Non-Spatial Queries)

Query Type	Correctly Resolved (%)	Partially Resolved (%)	Unresolved (%)
Spatial	82%	5%	13%
Non-Spatial	95%	4%	1%

Source: Source (2025)

3. *Dynamic Context Awareness*

By maintaining a short-term dialogue memory within sessions, the system was capable of building contextual understanding over time (Porcu, 2024). For instance, after a user mentioned "Colombo 5" and later asked for "another apartment near there," the system retained the context and adapted subsequent results accordingly. This conversational memory helped emulate a natural, human-like property search experience.

4. *Limitations and Future Enhancements*

Despite strong performance in most urban and semi-urban contexts, the system's effectiveness was limited in areas with inconsistent or sparse listing data. Since many property ads include only high-level location metadata (e.g., just the city or a vague street name), precise spatial filtering was occasionally compromised. This constraint highlighted the need for improved data structuring at the point of listing publication.

In future versions, the integration of satellite imagery, geotagging, and user behavior (clickstream) data will be explored to enhance spatial reasoning and personalize recommendations further. Additionally, multilingual support and colloquial language handling will be refined to better serve diverse user demographics.

Conclusion

In conclusion, this study demonstrates the significant potential of a spatially-aware conversational AI system to transform real estate platforms by enhancing query understanding and property relevance. By integrating NLP, geospatial intelligence, and user behavior analysis, the developed system effectively addresses the limitations of traditional chatbots in interpreting vague or spatially-constrained queries. The system's ability to automate query understanding, resolve spatial ambiguities, and maintain dynamic context awareness resulted in marked improvements in query interpretation and user experience. While data limitations in areas with sparse or inconsistent listing information present a challenge, future enhancements such as integrating satellite imagery, geotagging, and multilingual support promise to further refine the system's spatial reasoning and personalization capabilities. Ultimately, this approach paves the way for more intuitive, localized, and intelligent property discovery, benefiting both real estate platforms and users seeking to navigate complex and spatially-sensitive markets. (Bhaker, 2025).

Acknowledgement

The authors gratefully acknowledge **LankaPropertyWeb** for access to the data from their private database used in this study.

References

Bhaker, N. A. (2025). The technological transformation of real estate: How AI is reshaping the industry. *World Journal of Advanced Research and Reviews*, 26(2), 2315–2323. https://doi.org/10.30574/wjarr.2025.26.2.1826

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P. S., Yang, Q., & Xie, X. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 15(3), 1–45. https://doi.org/10.1145/3641289

Chellappan, R. B. (2025). From Scripted Responses to Contextual Understanding: A Comprehensive Examination of Chatbot Technology and Its Journey Toward Mimicking Human Conversations. . https://doi.org/10.2139/ssrn.5048174

Chen, H., Liu, X., Yin, D., & Tang, J. (2017). A Survey on Dialogue Systems: Recent Advances and New Frontiers. *SIGKDD Explorations Newsletter,* 19(2), 25–35. https://doi.org/10.48550/arXiv.1711.01731

Hayunga, D. K., & Pace, R. K. (2009). Spatial statistics applied to commercial real estate. *The Journal of Real Estate Finance and Economics,* 41(2), 103–125. https://doi.org/10.1007/s11146-009-9190-2

Hosen, B., Rahaman, M., Kumar, S., Sagar, L., & Akhtar, M. N. (2023). LEVERAGING ARTIFICIAL INTELLIGENCE AND BIG DATA FOR ADVANCED SPATIAL ANALYTICS AND DECISION

SUPPORT SYSTEMS IN GEOGRAPHY. 1(2), 62–67. https://doi.org/10.26480/magg.02.2023.62.67

Huizenga, E., & Yang, B. (2025). Agent Development Kit: Making it easy to build multi-agent applications. https://developers.googleblog.com/en/agent-development-kit-easy-to-build-multi-agent-applications/

Porcu, V. (2024). The Role of Memory in LLMs: Persistent Context for Smarter Conversations. *International Journal of Scientific Research and Management (IJSRM),* 12(11), 1673–1691. https://doi.org/10.18535/ijsrm/v12i11.ec04

Takyar, A., & Takyar, A. (2023, May 13). AI use cases in the real estate industry. LeewayHertz - AI Development Company. https://www.leewayhertz.com/ai-in-real-estate/

Vasdev, K. (2024). Integrating GIS with Natural Language Processing for Location-Based Insights.

International Journal for Multidisciplinary Research, 6(2). https://doi.org/10.36948/ijfmr.2024.v06i02.23437