# Urban Fusion: Geo Agents for Participatory Urban Intelligence

**Mingyang Sun**
Massachusetts Institute of Technology
msun14@mit.edu

## Abstract

Urban site selection is a complex task requiring the integration of diverse data sources including geospatial, demographic, and visual information. Traditional approaches using Geographic Information Systems (GIS) and rule-based methods struggle with unstructured data and real-world complexity. While recent advances in large language models (LLMs) and vision-language models (VLMs) have improved multimodal data processing, there exists no standardized benchmark to evaluate their effectiveness in site selection tasks. This paper introduces a novel benchmark for LLM-based site selection in Dubai, testing both single and multi-constraint problems across various modalities. I propose UrbanFusion, a multimodal geo-agent framework that leverages LLMs to process diverse data types (GIS, images, csv data) and employs reasoning capabilities through tool usage and code generation. Our initial experiments demonstrate the framework's ability to satisfy complex constraints and validate intermediate steps, establishing a foundation for more adaptive and explainable approaches to urban site selection. The GitHub repo is available here: Link.

## 1 Introduction

Urban site selection is a critical challenge in urban planning, requiring the integration of geospatial, socioeconomic, and qualitative data to make informed decisions. In rapidly developing cities like Dubai, selecting optimal locations for restaurants commercial establishments demands balancing multiple factors including proximity to customers, delivery efficiency, rental costs, and regulatory compliance. Traditional approaches such as Geographic Information Systems (GIS) have long focused on static metrics like proximity and cost-efficiency [Church, 2002, Aboulola, 2018]. While effective in controlled scenarios, these methods struggle to incorporate unstructured data (e.g., social sentiment, satellite imagery) or model dynamic stakeholder interactions [Batty, 2013].

Multi-agent systems (MAS) emerged to address these limitations, simulating urban dynamics through autonomous agents [Hosseinali et al., 2013, Ligtenberg et al., 2001]. However, traditional MAS frameworks lack the computational power to process multimodal inputs or resolve complex, competing constraints. Recent advancements in large language models (LLMs) and vision-language models (VLMs) have enabled automated analysis of textual and visual data [Achiam et al., 2023, Touvron et al., 2023, Feng et al., 2024], but these tools often operate in isolation, leaving a critical gap in holistic, context-aware urban decision-making.

In this paper, I introduce UrbanFusion, a multimodal geo-agent framework designed specifically for restaurant site selection in Dubai. Our first contribution is a novel benchmark for evaluating LLM-based site selection systems—the first of its kind—testing both single and multi-constraint problems across various modalities. The benchmark measures metrics such as accuracy and pass rate across different constraint categories, providing a standardized way to evaluate site selection models. Our second contribution is the UrbanFusion framework itself, which features: (1) multimodal fusion of geospatial poi data, consumer sale data, satellite imagery; (2) Reasoning ability of LLM on

agents' ability to use tools and write code to complete tasks. dynamic coordination via an LLM-based orchestrator to balance domain-specific agents;

## 2 Related Work

### 2.1 Traditional Site Selection Methods

Early GIS-based methods formalized spatial optimization through models like the Maximal Coverage Location Problem, prioritizing proximity to demand points [Church, 2002]. Aboulola [2018] extended these principles to small retail facilities, incorporating basic demographic variables but overlooking qualitative factors like consumer sentiment. While foundational, these approaches treated urban systems as static, failing to account for real-world complexities such as shifting zoning laws or evolving consumer behavior.

### 2.2 Multi-Agent Systems in Urban Planning

Agent-based frameworks emerged to model stakeholder dynamics in urban ecosystems. Hosseinali et al. [2013] simulated land-use changes in Qazvin, Iran, using rule-based agents to represent residents and businesses, while Ligtenberg et al. [2001] enforced zoning compliance through agent negotiations. However, these systems relied on predefined rules and lacked tools to process unstructured data or scale to multimodal urban environments.

### 2.3 Multimodal Data Integration

Recent work has explored integrating geospatial, textual, and visual data for urban analytics. Gao et al. [2022] demonstrated the effectiveness of pre-training language models on Points-of-Interest (POIs) to infer spatial patterns and urban functionality. Building on this, Chen et al. [2024] developed an LLM agent capable of fusing satellite imagery with textual metadata to automate geospatial analysis, addressing limitations in traditional parallel-input approaches. Despite progress, these studies treated modalities in isolation, missing opportunities for cross-domain synthesis.

### 2.4 LLM-Driven Urban Analytics

The integration of large language models (LLMs) like GPT-4 [Achiam et al., 2023] has advanced urban analytics by enabling structured interpretation of multimodal data and dynamic task decomposition. For example, UrbanLLM [Jiang et al., 2024] automates urban activity planning by decomposing natural language queries into executable workflows that coordinate domain-specific models. Similarly, CityGPT [Feng et al., 2024] enhances spatial reasoning in LLMs by embedding geospatial knowledge and fine-tuning models on tasks like location recommendation. However, existing frameworks primarily focus on isolated tasks, such as traffic prediction or facility siting, and lack robust mechanisms to resolve conflicts between competing objectives in multi-stakeholder urban environments.

To the author's knowledge, this paper is the first to introduce a comprehensive benchmark for evaluating LLM-based site selection systems. Unlike previous work that focuses on either traditional GIS methods or general-purpose LLM applications, this approach specifically addresses the challenges of multimodal data integration and constraint satisfaction in the context of urban site selection in Dubai.

## 3 Problem Statement

I formalize the urban site selection problem as a constraint satisfaction problem with multiple objectives. Let $S = \{s_1, s_2, \ldots, s_n\}$ be the set of all potential sites in the urban area of Dubai. Each site $s_i$ is characterized by a set of attributes $A_i = \{a_{i1}, a_{i2}, \ldots, a_{im}\}$ representing features such as location coordinates, area size, rental cost, and proximity to various points of interest.

The site selection problem involves identifying a subset of sites $S^* \subset S$ that satisfy a set of constraints $C = \{c_1, c_2, \ldots, c_k\}$ while optimizing one or more objective functions. Formally, I define:

$$S^* = \{s \in S \mid \forall\, c \in C,\ c(s) = \text{true}\}$$

For each constraint $c_j$, I define a satisfaction function:

$$c_j(s_i) = \begin{cases} \text{true} & \text{if constraint } j \text{ is satisfied by site } i, \\ \text{false} & \text{otherwise.} \end{cases}$$

I further define an objective function $f(s)$ that evaluates the quality of a site based on relevant metrics. The optimization problem becomes:

$$\text{maximize } f(s) \quad \text{subject to} \quad \forall\, c \in C,\ c(s) = \text{true}.$$

For evaluation, I define the accuracy rate (AR) and recall rate (RR) against ground truth lists:

$$\text{AR} = \frac{|S^* \cap GT|}{|S^*|}, \quad \text{RR} = \frac{|S^* \cap GT|}{|GT|}$$

where $GT$ represents the ground truth set of optimal sites.

# 4 Proposed Approach

## 4.1 Benchmark Development

The benchmark for evaluating LLM-based site selection systems serves as my first major contribution, as no standardized evaluation framework currently exists for this domain. This benchmark is structured to test both the multimodal data integration capabilities and the reasoning abilities of LLM-based systems.

Each problem set contains both single-constraint and multi-constraint problems to evaluate how well systems handle individual constraints and manage competing demands. The benchmark incorporates six primary constraint categories that reflect the multifaceted nature of urban site selection in Dubai:

- **Geospatial Constraints**: Proximity (e.g., 1-3 km delivery radius), zoning (e.g., commercial land use only)
- **Temporal Constraints**: Peak hours (e.g., lunch rush), seasonality (e.g., Ramadan special hours)
- **Economic Constraints**: Cost (e.g., rent limits), revenue (e.g., daily GMV), ROI (e.g., annual ROI)
- **Market Constraints**: Competition (e.g., same cuisine density), demand (e.g., orders/day)
- **Demographic Constraints**: Population density, age group, income level
- **Operational Constraints**: Logistics (e.g., kitchen size), compliance (e.g., Halal certification)

The primary evaluation metrics include pass rate, constraint satisfaction rate by category, resolution time, and explanation quality.

## 4.2 Multimodal Embedding Database

To support the multimodal nature of site selection, I implement a multimodal embedding database that integrates diverse data types through a hierarchical fusion approach:

$$E_{\text{fused}} = F(E_{\text{GIS}}, E_{\text{CSV}}, E_{\text{image}}, E_{\text{doc}})$$

where $E$ represents embeddings of different modalities and $F$ is a fusion function defined as:

$$F(E_1, E_2, \ldots, E_n) = W_1 \cdot E_1 + W_2 \cdot E_2 + \cdots + W_n \cdot E_n,$$

with learned weights $W_i$ that determine the contribution of each modality to the final representation.

I implement this using a combination of pre-trained models:**For GIS data:** I use a custom encoder that converts geospatial features into dense vectors. **For CSV data:** I apply a tabular encoder that

captures numerical and categorical relationships. **For image data:** I utilize a vision transformer to extract features from satellite imagery.

These embeddings are stored in a vector database (Chroma) that supports efficient similarity search across modalities using:

$$\text{sim}(q, d) = \cos\big(E_{\text{fused}}(q), E_{\text{fused}}(d)\big),$$

where $q$ represents a query and $d$ represents a document in the database.

### 4.3 Agent Framework Architecture

The UrbanFusion framework is implemented using LangChain and consists of multiple specialized agents coordinated by a central orchestrator:

1. **Coordinator Agent**: The central orchestration component responsible for managing agent registration, invocation, and collaboration.
2. **Specialized Agents**:
   - **GeoAgent**: Handles geospatial queries and analysis
   - **Constraints Agent**: Enforces and evaluates both hard constraints and soft preferences
   - **Evaluation Agent**: Validates results against established benchmarks
   - **Explanation Agent**: Generates interpretable explanations of site recommendations
3. **Functional Tools**:
   - **MapTool**: Generates interactive maps and visualizes geospatial data
   - **Data Analysis Tool**: Processes structured data like financial metrics
   - **Visualization Tool**: Creates charts and graphs to support decision-making
   - **Code Generation Tool**: Enables agents to write and execute Python code for complex geospatial analysis

The agent selection and execution process follows:

1. Parse user query $q$ into intent $i$ and constraints $C$.
2. Select appropriate agents $A = \{a_1, a_2, \ldots, a_n\}$ based on intent.
3. Execute agents in parallel or sequence based on dependencies.
4. Aggregate results using weighted voting: $R = \sum_i w_i \cdot r_i$
5. Generate explanation $E$ based on results $R$ and process trace.

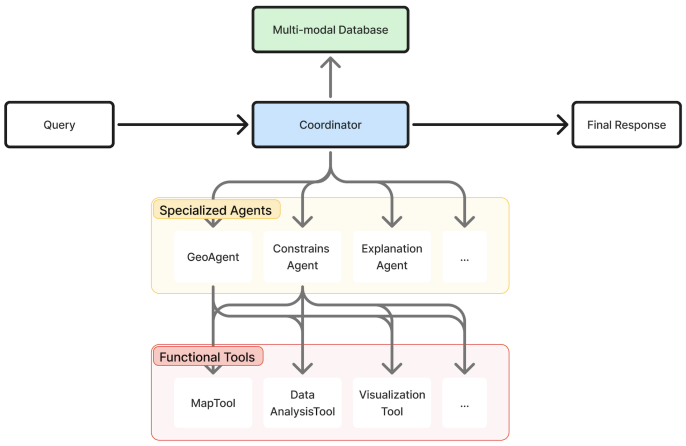


Figure 1: UrbanFusion framework integrates multimodal data sources through specialized agents to provide constraint-aware site recommendations.

# 5 Experimental Methodology

## 5.1 Dataset Description

The UrbanFusion prototype works with the following data sources.

- **CSV Data:** The file `talabat_sample.csv` contains fields including `vendor_location` (POINT coordinates), `customer_location` (POINT coordinates), `order_timestamp` (datetime), `cuisine_type` (categorical), `delivery_time` (calculated from timestamps), `gmv_amount` (financial metric), `delivery_fee` (cost factor), and `order_volume` (count per location).
- **GIS Data:** OpenStreetMap (OSM) data, downloaded from OpenStreetMap, provides spatial information including boundaries, road networks, and existing facilities.
- **Image Data:** Satellite images offer visual representations of urban patterns.

## 5.2 Preliminary Implementation

The prototype implementation integrates a knowledge base to organize and retrieve multimodal data. This implementation focuses on demonstrating the core components of the UrbanFusion framework without training new models. The prototype uses LangChain to design an agent that can interpret natural language queries and decompose them into tool-assisted subtasks. The agent employs the reasoning capabilities of a large language model (LLM) to decide which tools to use and in what sequence, thus creating a dynamic multimodal agent. Two key approaches are demonstrated:

- LLM with reasoning (using an LLM that can employ reasoning capabilities to leverage tools and generate code for complex tasks).
- LLM without reasoning (using an LLM in a more traditional question-answering format without the ability to use tools or generate code). This comparison allows for an evaluation of how reasoning capabilities impact the performance of LLMs in constraint-driven urban site selection tasks.

## 5.3 Evaluation

The evaluation methodology for the prototype focuses on two aspects: a final list that assesses accuracy and pass rate against benchmark problem sets, and intermediate tests based on step results provided by the agents. The evaluation does not involve training or fine-tuning models; rather, it assesses the capabilities of existing LLMs when applied to the urban site selection task with and without reasoning abilities. The prototype is designed to demonstrate the potential of the approach rather than to provide quantitative performance metrics, which would require the completed benchmark development described in the next steps section.

# 6 Results and Discussion

The UrbanFusion framework represents a significant advancement in urban site selection methodology by integrating multimodal data sources, leveraging LLM capabilities, and implementing a constraint-aware agent architecture. The initial prototype focuses on the Dubai use case, selecting restaurant locations that align with delivery demands and consumer preferences while minimizing emissions.

## 6.1 Benchmark Performance

The prototype demonstrates the potential for evaluating the UrbanFusion framework against the benchmark for site selection in Dubai, testing its ability to handle both single and multi-constraint problems across different complexity levels. The implementation allows for comparing LLMs with reasoning abilities against those without reasoning capabilities.

LLMs with reasoning abilities can interpret natural language queries about site selection, break down complex queries into manageable subtasks, use tools to access and process multimodal data, generate code to perform spatial analysis, and provide explanations for recommendations. In contrast, LLMs

without reasoning abilities are limited to answering questions based on pre-existing knowledge, processing text inputs without accessing external tools, providing general recommendations without spatial analysis, and lacking the ability to verify constraints through computation.

Thus, the prototype implementation demonstrates the framework's architecture and components rather than offering quantitative performance metrics, which would require the completed benchmark development described in the next steps.

### 6.2 Multimodal Reasoning Analysis

The prototype also demonstrates the potential value of multimodal integration in urban site selection. The architecture shows how different data types can be incorporated: text data provides context about requirements and constraints; geospatial data enables location-based analysis and proximity calculations; and visual data from satellite imagery informs about urban patterns and visibility. The combination of text and geospatial data is particularly important for site selection problems, reflecting their inherently spatial nature. Moreover, adding visual information from satellite imagery has the potential to enhance performance for constraints related to visibility and urban patterns.

The prototype further demonstrates the potential of code generation capabilities for geospatial analysis, as the architecture includes a Code Generation Tool that enables agents to write and execute Python code for complex spatial analyses beyond the capabilities of standard LLMs.

## 7 Next Steps and New Research Ideas

### 7.1 Next Steps

The immediate next steps for the UrbanFusion project include:

- Finish benchmark development with clearly defined problem sets of varying complexity.
- Establish clear goals for test accuracy and pass rate to compare different LLM performances.
- Complete the implementation of all components in the agent framework with all tools.
- Expand the dataset with more comprehensive GIS and image data.

These steps are essential to move from the current prototype to a fully functional system that can be rigorously evaluated against the benchmark.

### 7.2 Future Research Directions

Looking beyond the immediate next steps, several promising research directions emerge:

- **Enhanced Multimodal Fusion:** Although the current prototype demonstrates the architecture for multimodal fusion, future work could explore additional integration techniques for combining geospatial, temporal, and visual data in urban site selection.
- **Advanced Constraint Resolution:** Develop more sophisticated mechanisms for resolving competing constraints in complex urban environments, particularly when temporal and economic factors conflict with geospatial recommendations.
- **Cross-City Transfer Learning:** Investigate how knowledge and models developed for Dubai could be transferred to other urban contexts with different characteristics and constraints.
- **Domain Extension:** Apply the UrbanFusion approach to broader urban planning domains such as affordable housing placement, public facility planning, and sustainable urban development.

The long-term vision is to create a flexible, general-purpose tool that urban planners, policymakers, and developers can use across various planning scenarios. By offering transparent, data-driven, and constraint-aware recommendations, UrbanFusion aims to promote more inclusive, efficient, and sustainable urban development.

## References


O. I. Aboulola. Gis spatial analysis: A new approach to site selection and decision making for small retail facilities. Master's thesis, The Claremont Graduate University, 2018.

J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, et al. Gpt-4 technical report. preprint arXiv:2303.08774, arXiv, 2023.

M. Batty. *The New Science of Cities*. MIT Press, 2013.

Y. Chen, W. Wang, S. Lobry, and C. Kurtz. An llm agent for automatic geospatial data analysis. preprint arXiv:2410.18792, arXiv, 2024. URL https://doi.org/10.48550/arXiv.2410.18792.

R. L. Church. Geographical information systems and location science. *Computers & Operations Research*, 29(6):541–562, 2002. doi: 10.1016/S0305-0548(99)00104-5. URL https://doi.org/10.1016/S0305-0548(99)00104-5.

J. Feng, Y. Du, T. Liu, S. Guo, Y. Lin, and Y. Li. Citygpt: Empowering urban spatial cognition of large language models. preprint arXiv:2406.13948, arXiv, 2024. URL https://doi.org/10.48550/arXiv.2406.13948.

Y. Gao, Y. Xiong, S. Wang, and H. Wang. Geobert: Pre-training geospatial representation learning on point-of-interest. *Applied Sciences*, 12(24):12942, 2022.

F. Hosseinali, A. A. Alesheikh, and F. Nourian. Agent-based modeling of urban land-use development: Case study: Simulating future scenarios of qazvin city. *Cities*, 31:105–113, 2013.

Y. Jiang, Q. Chao, Y. Chen, X. Li, S. Liu, and G. Cong. Urbanllm: Autonomous urban activity planning and management with large language models. preprint arXiv:2406.12360, arXiv, 2024. URL https://doi.org/10.48550/arXiv.2406.12360.

A. Ligtenberg, A. K. Bregt, and R. Van Lammeren. Multi-actor-based land use modelling: Spatial planning using agents. *Landscape and Urban Planning*, 56(1-2):21–33, 2001.

H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, others, and T. Scialom. Llama 2: Open foundation and fine-tuned chat models. preprint arXiv:2307.09288, arXiv, 2023.


## A Appendix

### A.1 Constraint Categories and Sample Questions

This appendix provides the constraint categories used in the UrbanFusion framework, followed by sample benchmark questions that incorporate these constraints.

The UrbanFusion framework addresses six primary constraint categories that reflect the multifaceted nature of urban site selection in Dubai, see Table 1.

The following sample questions demonstrate how the UrbanFusion framework handles queries with varying constraints:

1. "Find suitable locations for a new Italian restaurant within 3 km of Downtown Dubai, in a commercial zone only, with minimum 5,000 AED daily GMV potential."
2. "Identify optimal locations for a new fast-food restaurant within 2 km of residential areas, with maximum 3 similar cuisine restaurants within 1 km, less than 15% delivery cost ratio, at least 100 orders/day potential, and within 500m of arterial roads."
3. "Recommend locations for a premium dining restaurant with green zone adjacency for outdoor seating, targeting customers with 25k+ AED/month income, minimum 20% annual ROI, optimized for dinner peak (19:00-21:00), minimum 20m street frontage, halal certification compliance, at least 50m² kitchen size, in an expat-friendly zone, with no more than 3 same cuisine restaurants within 1km."

Table 1: UrbanFusion Site Selection Constraints

| Category | Subcategory | Option | Data Requirement |
|---|---|---|---|
| Geospatial | Proximity | 1-3 km delivery radius | Customer/Vendor coordinates |
| Geospatial | Proximity | $\leq$500m to arterial roads | OSM road network |
| Geospatial | Zoning | Commercial land use only | Dubai municipality GIS |
| Geospatial | Zoning | Green zone adjacency | Sentinel-2 imagery |
| Geospatial | Visibility | $\geq$20m street frontage | Google Street View |
| Temporal | Peak Hours | 12:00-14:00 lunch rush | Order timestamps |
| Temporal | Peak Hours | 19:00-21:00 dinner peak | Vendor prep times |
| Temporal | Seasonality | Summer demand spikes | Historical GMV patterns |
| Temporal | Seasonality | Ramadan special hours | Cultural calendar |
| Economic | Cost | $\leq$250 AED/m$^2$ rent | Bayut/Dubizzle data |
| Economic | Cost | $\leq$15% delivery cost | Route distance matrix |
| Economic | Revenue | req 5K AED daily GMV | Order basket values |
| Economic | ROI | req 20% annual ROI | Financial records |
| Market | Competition | $\leq$3 same cuisine/1km | Vendor locations |
| Market | Competition | $\geq$500m from major chains | Business registry data |
| Market | Demand | $\geq$100 orders/day | Order density heatmaps |
| Market | Demand | Untapped area potential | Population vs coverage |
| Demographic | Population Density | $\geq$5k/km$^2$ residents | Census data |
| Demographic | Age Group | 20-35yo majority | Survey data |
| Demographic | Income Level | $\geq$25k AED/month | Tax records |
| Demographic | Cultural | Expat-friendly zone | Migration patterns |
| Operational | Logistics | $\geq$50m$^2$ kitchen size | Satellite imagery |
| Operational | Logistics | 24/7 delivery access | Road closure data |
| Operational | Compliance | Halal certification | Regulatory docs |

4. "Find locations for a cloud kitchen with 24/7 delivery access, within 500m of arterial roads, maximum 15% delivery cost, and minimum 100 orders/day potential."
5. "Suggest locations for a seasonal pop-up restaurant that can capitalize on summer demand spikes, with green zone adjacency, targeting the 20-35 age demographic, and in areas with untapped market potential."

### A.2 Query Processing Workflow Example

This section demonstrates the workflow of the UrbanFusion framework when processing a site selection query.

**User Query:**
"Find suitable locations for a new Italian restaurant within 3 km of Downtown Dubai, in a commercial zone only, with minimum 5,000 AED daily GMV potential."

**Workflow:**

```
Step 1: Constraint Identification and Classification
- Geospatial/Proximity: Within 3 km of Downtown Dubai
- Geospatial/Zoning: Commercial zone only
- Economic/Revenue: Minimum 5,000 AED daily GMV potential

Step 2: Data Retrieval
- Access customer/vendor coordinates database for Downtown Dubai reference
- Access Dubai municipality GIS data for commercial zoning information
- Access order basket values data for GMV estimation

Step 3: Constraint Processing
- GeoAgent: Create 3 km buffer around Downtown Dubai coordinates
- GeoAgent: Intersect buffer with commercial zoning layer
```

```
- GeoAgent: Identify all potential commercial locations within buffer

Step 4: Economic Analysis
- For each potential location, calculate estimated daily GMV based on:
  * Nearby order density from talabat_sample.csv
  * Average order values in area
  * Similar cuisine performance in proximity

Step 5: Constraint Satisfaction Verification
- Constraints Agent: Verify each location against all constraints
- Constraints Agent: Filter locations with estimated GMV > 5,000 AED

Step 6: Explanation Generation
- Explanation Agent: Generate detailed rationale for each recommendation
- Explanation Agent: Include supporting data points and visualization
```

**Sample Output:**
"Based on your requirements, I've identified 5 potential locations for your Italian restaurant:

1. Location A: Business Bay Area (coordinates: 25.186, 55.282)
   - 1.8 km from Downtown Dubai
   - Zoned commercial (Dubai Municipality code: C-1)
   - Estimated daily GMV: 7,200 AED based on 48 similar orders/day at 150 AED average
2. Location B: DIFC North Gate (coordinates: 25.211, 55.275)
   - 2.1 km from Downtown Dubai
   - Zoned commercial (Dubai Municipality code: C-3)
   - Estimated daily GMV: 6,800 AED based on 42 similar orders/day at 162 AED average

[Additional locations would be listed with similar detailed analysis]"

### A.3 Evaluation Metrics

Table 2: Detailed Evaluation Metrics and Scoring System

| **Metric** | **Scale** | **Calculation Method** |
|---|---|---|
| Accuracy Rate (AR) | 0-100% | $\frac{\lvert S^* \cap GT \rvert}{\lvert S^* \rvert}$ where $S^*$ is the recommended set and $GT$ is the ground truth set |
| Recall Rate (RR) | 0-100% | $\frac{\lvert S^* \cap GT \rvert}{\lvert GT \rvert}$ where $S^*$ is the recommended set and $GT$ is the ground truth set |
| Resolution Time | Seconds | Time from query submission to final response |
| Tool Usage Effectiveness | 1-5 scale | Expert evaluation of appropriate tool selection and usage |
| Code Generation Quality | 1-5 scale | Correctness, efficiency, and readability of generated code |
| Explanation Completeness | 1-5 scale | Coverage of all relevant factors in explanations |
| Explanation Transparency | 1-5 scale | Clarity in explaining decision rationale |
| Explanation Actionability | 1-5 scale | Usefulness of explanations for decision-making |

The primary evaluation approach compares the recommended locations against ground truth lists rather than attempting to measure individual constraint satisfaction. This approach aligns with the mathematical formulation in the paper:

$$\text{AR} = \frac{|S^* \cap GT|}{|S^*|}, \quad \text{RR} = \frac{|S^* \cap GT|}{|GT|}$$

where $S^*$ represents the set of recommended sites and $GT$ represents the ground truth set of optimal sites.