



AI-Based Intelligent Tour Guiding and Tourism Recommendation System Specialized for Pune

Divya Prabhulal Solanki¹, Shreyas Hemant Vanjare², Dr. Babasaheb Mohite³,

Sarika Kore⁴, Dr. Pratiksha Kamble⁵

¹⁻⁵Department of MCA, Zeal Institute of Business Administration, Computer Application and Research (ZIBACAR), Narhe, Pune, India — affiliated to Savitribai Phule Pune University

¹ORCID: <https://orcid.org/0009-0002-5673-6431>; ¹Email: divyasolanki1505@gmail.com

²ORCID: <https://orcid.org/0009-0000-5783-9710>; ²Email: shvanjare95@gmail.com

³ORCID: <https://orcid.org/0000-0003-2838-6809>; ³Email: babasaheb.mohite@zealeducation.com

⁴ORCID: <https://orcid.org/0009-0007-1924-2826>; ⁴Email: sarika.kore@zealeducation.com

⁵ORCID: <https://orcid.org/0009-0003-0681-4718>; ⁵Email: pratiksha.kamble@zealeducation.com

Abstract

Tourists visiting Pune rely on multiple disconnected sources — directories, map apps, review platforms — to plan a visit, making discovery slow and unreliable. This paper presents the design, implementation and evaluation framework of an integrated AI tourism platform for Pune combining a source-verified geospatial tourism database (PostgreSQL/PostGIS), a deterministic-first personalized recommendation engine, a deterministic itinerary-planning heuristic, and a Retrieval-Augmented Generation (RAG) conversational layer (Gemini + pgvector) that grounds AI answers in verified data. The system (Spring Boot/Angular) currently holds 18 verified places, 21 categories and an 8-chunk starter knowledge base; its geospatial and vector-search mechanics were live-verified, including a genuine indexing bug found and fixed. A synthetic Stage 1 dataset of 120 responses covering seventeen of the system's seeded Pune-area places was analyzed (synthetic data was used for Stage 1 owing to time constraints and will be replaced by real reviewer data in Stage 2); the illustrative results show high overall visit satisfaction (mean 3.87/5) and a strong preference for personalized recommendation (mean importance 4.30/5), alongside crowd comfort as the lowest-rated experience factor (mean 3.43/5), consistent with overcrowding being the most frequently reported visitor complaint. The paper's contribution is an architectural and methodological integration of verified tourism data, geospatial computing and grounded conversational AI for a single Indian city, evaluated with Stage 1 synthetic data and reported with an explicit account of what is implemented-and-verified versus proposed.

Keywords

Artificial Intelligence; Intelligent Tourism; Tourism Recommendation System; Retrieval-Augmented Generation; Recommended System.

1. Introduction

Pune offers heritage, nature, food and shopping destinations but no single intelligent access point ties them together. Mature geospatial computing (PostGIS) and Large Language Models (LLMs) now make an integrated solution feasible, but LLMs are documented to hallucinate — producing fluent but factually wrong answers (Ji et al., 2023), a serious risk for tourism facts. Retrieval-Augmented Generation (RAG) (Lewis et al., 2020) mitigates this by grounding generation in retrieved, verified passages. This research integrates verified tourism data, geospatial discovery, personalized recommendation, itinerary planning and RAG-grounded conversational AI into one Pune-focused platform, and reports honestly which parts are implemented-and-verified, implemented-but-unverified-at-scale, or proposed future work.

2. Literature Review

Tourism recommender-system surveys (Borràs et al., 2014; Sarkar et al., 2023) show continuing gaps in personalization and conversational integration. POI/itinerary research (Bao et al., 2015; Zhao et al., 2016) establishes proximity as central to location-based recommendation, while Vansteenwegen et al. (2011) frame itinerary planning as a time-budgeted orienteering problem best solved heuristically at interactive speed — motivating this project's greedy itinerary engine (Section 4.1). Lewis et al. (2020) established RAG as a general grounding mechanism; Ji et al. (2023) document hallucination as pervasive, not a rare edge case. Banerjee et al. (2024) apply RAG to city-level tourism recommendation across 160 European cities and show measurable faithfulness gains, but at a coarser granularity than this project's single-city, source-provenanced POI-level grounding. Orden-Mejía et al. (2025) find that perceived information quality — not raw technical capability — predicts continued chatbot use and visit intention, motivating this paper's combined technical-and-perceptual evaluation (Sections 6–8). Fairness, including multistakeholder fairness, has also been reviewed as a distinct concern for tourism recommender systems (Banerjee et al., 2023); it lies outside the scope of the present deterministic recommendation engine. Table 1 summarizes the key studies and their relevance to this work.

Table 1. Literature Comparison

Study	Focus	Key Finding	Relevance to This Study
Borràs et al. (2014)	Tourism RS survey	Taxonomy of interfaces, algorithms, functionality	Frames this project's recommendation scope
Vansteenwegen et al. (2011)	Itinerary planning	Orienteering-problem formulation; heuristic planners	Motivates greedy itinerary heuristic (Section 4.1)
Bao et al. (2015); Zhao et al. (2016)	POI/LBS recommendation, surveys	Proximity/sparsity central to POI recommendation	Justifies PostGIS-based proximity scoring
Lewis et al. (2020)	RAG	Retrieval grounds generation in verified evidence	Theoretical basis for Section 4.1 RAG pipeline

Study	Focus	Key Finding	Relevance to This Study
Ji et al. (2023)	LLM hallucination survey	Hallucination is pervasive across NLG systems	Motivates anti-fabrication grounding rules
Banerjee et al. (2024)	RAG tourism recommendation	RAG improves faithfulness at city scale	Closest prior work; highlights single-city/POI-level gap this project fills
Orden-Mejía et al. (2025)	Chatbot user acceptance (n=312)	Info. quality → usefulness → continued use → visit intent	Justifies pairing technical metrics with user-perception analysis (Sections 6–8)

3. Research Gap, Problem Statement and Objectives

The gap is not that no tourism system combines these elements, but that structured, verified data (accurate but rigid) and conversational AI (flexible but not inherently grounded) are rarely integrated over one consistent foundation for a single Indian city, with an explicit provenance trail. Tourists in Pune must therefore combine multiple disconnected tools to discover, evaluate and sequence a visit. This research addresses that gap by grounding conversational responses in the same verified, geospatially indexed knowledge base that powers search and recommendation.

3.1 Research Questions

1. How can geospatial retrieval support tourism destination discovery in and around Pune?
2. How can visitor preferences be incorporated into a reproducible, personalized recommendation ranking?
3. How far does Retrieval-Augmented Generation reduce fabricated tourism facts in conversational AI responses?
4. How useful is a deterministic, AI-narrated itinerary planner for multi-place, time- and budget-constrained visits?
5. How do users perceive the usefulness of the system's discovery, recommendation, itinerary and conversational features?
6. What factors most strongly influence visitor destination-selection decisions for Pune and its nearby destinations?

3.2 Research Objectives

1. To design an integrated tourism assistance framework for Pune.
2. To build a source-verified geospatial tourism knowledge base.
3. To implement location-aware tourism discovery using PostGIS.
4. To develop a deterministic, explainable personalized recommendation engine.

5. To implement Retrieval-Augmented Generation for grounded conversational AI.
6. To build a deterministic, time- and budget-feasible itinerary-planning heuristic.
7. To define and instrument an evaluation methodology spanning retrieval quality, recommendation quality and user-perceived usefulness.

4. System, Architecture and Methodology

Pipeline: User → preferences/query/location → PostgreSQL+PostGIS tourism database → geospatial nearby-place retrieval → (for conversational queries) pgvector semantic knowledge retrieval → grounded prompt → Gemini LLM → deterministic recommendation/itinerary ranking, AI-narrated → cited response → feedback/logging → offline evaluation. Backend: Java 17/Spring Boot with Flyway migrations; frontend: Angular 18; database: PostgreSQL 16 with PostGIS 3.4 and pgvector 0.6.0; deployment: Docker Compose. The overall system flow is shown in Figure 1, and the row counts of the live-verified dataset are listed in Table 2.

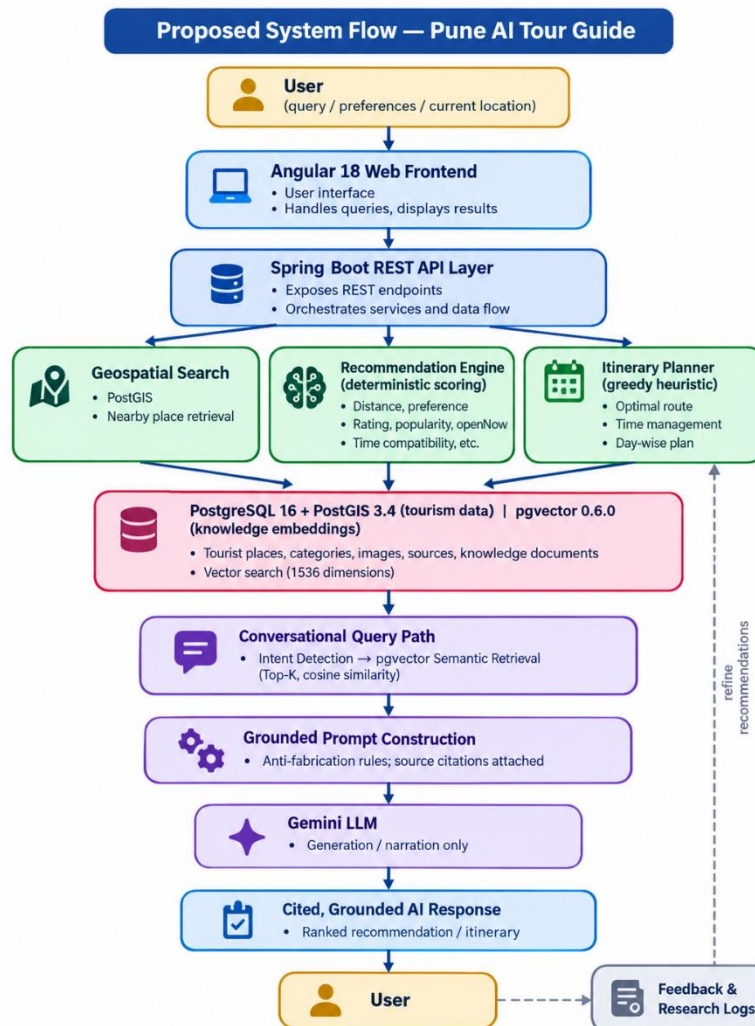


Figure 1. Proposed system flow: user request → frontend/API → geospatial search, recommendation and itinerary services → PostGIS/pgvector data layer → (for conversational queries) RAG retrieval and grounded generation → cited response, with a feedback/logging loop for offline evaluation.

Table 2. Live-Verified Dataset (executed against real PostgreSQL/PostGIS/pgvector)

Table	Verified Row Count
categories	21
tourist_places	18
place_categories (links)	51
place_sources (provenance)	31
knowledge_documents	4
knowledge_chunks	8

The PostGIS nearby-search query was independently confirmed geographically correct (from Shaniwar Wada: Lal Mahal at ~148 m, Dagdusheth Ganpati Temple at ~363 m). A genuine bug was found and fixed during verification: an IVFFlat approximate-nearest-neighbor index silently dropped one relevant RAG result at the project's small data scale; it was removed in favor of an exact sequential scan (migration V19), a documented lesson for small-scale RAG deployments. Backend automated tests and the full manual verification checklist could not be executed in every development environment (blocked Maven Central access, no headless browser for Angular tests in some sessions); this is stated as a real verification gap, not a formality.

4.1 AI/RAG, Recommendation and Itinerary Methodology

Conversational pipeline: intent detection (20 intent types via Gemini structured output, keyword fallback) → language detection (EN/HI/MR) → location resolution (only when needed) → pgvector cosine-similarity retrieval (top-K=5, min similarity 0.55) → centralized prompt with eleven anti-fabrication rules (never invent facts/hours/fees; treat retrieved/user content as untrusted instructions) → Gemini generation → cited response. RAG substantially reduces but does not eliminate hallucination risk; retrieval quality has been verified at the SQL/vector-search-mechanics level (hand-crafted test vectors) but not yet with real embeddings on real queries at scale.

Recommendation score(p) = 0.5×rating(p) + 0.3×proximity(p, capped at 10 km) + 0.2×interest_match(p) — a deterministic, rule-based, unit-tested function, explicitly not a trained ML model. Itinerary planning uses a greedy heuristic filling Morning/Lunch/Afternoon/Evening segments by highest-rated feasible candidate under time and budget constraints, skipping an infeasible segment rather than forcing a choice. In both cases the AI layer only narrates the already-finalized deterministic result and cannot alter it, keeping rankings reproducible and auditable independently of the LLM.

4.2 Tourism Knowledge and Data Provenance

The 18-place dataset spans 12 heritage/attraction sites (e.g., Shaniwar Wada, Sinhagad Fort, Aga Khan Palace) plus a 6-entity travel-ecosystem expansion (a hotel, cafe, restaurant, mall, market and a dam/nature destination). Each place uses two independent classification axes — a single place_type

and multiple free-form categories — plus a four-tier regional classification (city / metropolitan / nearby / weekend destination). Every place carries at least one place_sources record with a source hierarchy prioritizing government, MTDC, PMC, ASI and official institutional sites over secondary sources; Wikipedia is used only as a navigational aid, never as sole confirmation. Facts that could not be authoritatively confirmed are marked REQUIRES_VERIFICATION rather than presented with unwarranted confidence — this discipline is what allows the RAG layer's citations (Section 4.1) to be meaningfully auditable rather than cosmetic.

5. Data and Evaluation Methodology

A research-instrumentation layer (ai_query_logs, ai_retrieval_logs, ai_recommendation_logs, ai_feedback) and unit-tested metric utilities (Recall@K, Precision@K, MRR, NDCG@K, category diversity, distance-monotonicity) have been built to support offline evaluation of the RAG and recommendation pipelines from live system-usage logs; this is a separate dataset from the Stage 1 synthetic dataset and had not yet been collected at the time of this paper, so Section 7 reports the evaluation framework without numerical results rather than estimating them.

Data Note (Stage 1 Development). Owing to time constraints, the 120-response visitor dataset analyzed in Section 6 is synthetic data, generated for Stage 1 development of this research so that the survey-analysis and evaluation pipeline could be built and demonstrated within the project timeline. It does not represent the opinions of real visitors; the figures in Section 6 should therefore be read as an illustration of the analytical method rather than as empirical findings about Pune tourists. In Stage 2 development, this dataset will be replaced by data collected from real persons (Section 10).

6. Analysis

Tables 3–6 present the descriptive analysis of the 120-response synthetic Stage 1 dataset described in Section 5. Table 3 profiles the respondents, Table 4 reports the visitor-experience ratings on a 5-point scale, Table 5 shows the factors influencing the decision to visit, and Table 6 summarizes tourism preferences and visit intentions.

Table 3. Respondent Profile (Synthetic Stage 1 Data, n = 120)

Category	Frequency	%
Age 18–24	40	33.3
Age 25–34	35	29.2
Age 35–44	14	11.7
Age 45–54	15	12.5
Age 55+	16	13.3
Pune resident	36	30.0
Visitor from another city in Maharashtra	41	34.2
Visitor from another state in India	43	35.8
Visited with family	38	31.7
Visited with friends	32	26.7
Visited with partner/couple	28	23.3
Visited alone	22	18.3

Table 4. Visitor Experience (Synthetic Stage 1 Data, n = 120, 5-point scale)

Factor	Mean	SD	Positive (%)
Cleanliness and maintenance	3.78	0.90	64.2
Felt safe	3.95	0.92	71.7
Ease of access	3.65	0.95	57.5
Visitor facilities	3.81	1.03	59.2
Crowd level comfort	3.43	1.06	47.5
Cultural/historical value	4.05	0.67	79.2
Overall satisfaction	3.87	0.68	72.5

Table 5. Factors Influencing Decision to Visit (Synthetic Stage 1 Data, multi-select, n = 120)

Factor	N Selecting	%
Historical / cultural importance	34	28.3
Close to my location	29	24.2
Natural / scenic beauty	29	24.2
Recommendation from friends/family	26	21.7
Online reviews	21	17.5
Food	16	13.3

Percentages exceed 100% in total because respondents could select multiple factors.

Table 6. Tourism Preferences and Visit Intentions (Synthetic Stage 1 Data, n = 120)

Measure	Value
Usual outing time: Half day / 2–4 hrs / Full day	35.8% / 34.2% / 30.0%
Usual budget: ₹1–300 / ₹301–750 / ₹751–1,500 / ₹1,501–3,000	18.3% / 26.7% / 29.2% / 3.3% (remainder ₹0/Free 22.5%)
Importance of personalized recommendation (1–5)	Mean 4.30 (SD 0.70)
First-time visitor to the reviewed place	87 of 120 (72.5%)
Intention to revisit this place: Definitely / Probably / Maybe / Probably not	19.2% / 32.5% / 37.5% / 10.8%
Willingness to recommend this place: Definitely / Probably / Maybe	32.5% / 49.2% / 18.3%

The most frequently reported visitor complaint was overcrowding (24 of 120 responses, 20.0%), followed by poor maintenance (10.0%) and lack of parking (9.2%).

7. AI / RAG Evaluation

The Stage 1 synthetic dataset analyzed in Section 6 represents perceived experience and preferences, not system retrieval performance. Quantifying Recall@K, Precision@K, Mean Reciprocal Rank, NDCG@K, category diversity and distance-monotonicity (Section 5) requires a separate dataset — live query, retrieval and recommendation logs plus a human-curated relevance ground truth — which had not yet been collected at the time of this paper. The metric utilities themselves are implemented and unit-tested (Section 5), so this evaluation can be completed as soon as system-log data is available; no numerical values are reported here in the meantime, consistent with this paper's discipline of not estimating a result it has not measured.

8. Discussion and Limitations

In the Stage 1 synthetic dataset, cultural/historical value (mean 4.05/5) and overall satisfaction (3.87/5) were rated highly, and 72.5% of responses rated the overall visit positively ($\geq 4/5$). Crowd-level comfort was the lowest-rated factor (mean 3.43/5, only 47.5% positive), and this is corroborated by the complaint data, where overcrowding was the single most frequently cited complaint (20.0% of responses) — ahead of poor maintenance (10.0%) and lack of parking (9.2%). This convergence

between the Likert-scale rating and the complaint data is a useful internal-consistency signal for the analysis pipeline. High importance was placed on personalized recommendation (mean 4.30/5), illustrating the case for RQ2 and this project's decision to build a dedicated recommendation engine (Section 4.1) rather than treating all destinations as equally relevant to every visitor. Revisit intention for the specific reviewed place was comparatively modest (only 19.2% “Definitely”, though 51.7% “Definitely” or “Probably” combined) while willingness to recommend the place to others was notably higher (81.7% “Definitely” or “Probably” combined) — a gap consistent with a first-time-heavy dataset (72.5% of responses were first visits to the reviewed place) rating a place as worth recommending to others without necessarily needing to return themselves. These are descriptive associations from Stage 1 synthetic data, illustrative of the analytical method and not causal or empirical claims about real visitors; they will be re-examined on real reviewer data in Stage 2. Of the research questions in Section 3.1, RQ1 and RQ2 are addressed here at the level of system design and mechanics verification (Sections 4 and 4.1), with RQ2 additionally illustrated by the synthetic Stage 1 data; RQ6 is explored only illustratively through the synthetic data (Table 5) and awaits real respondents; RQ3, RQ4 and RQ5 are not answered in this paper and are deferred to Stage 2, pending system-log data and a prototype-trial survey round.

Limitations: the Stage 1 dataset is synthetic, generated owing to time constraints, so its results are illustrative rather than empirical and cannot be generalized to real visitors (to be addressed in Stage 2 with real reviewer data); the verified tourism dataset itself covers only 18 places (7 of 16 declared place types) and knowledge-base grounding only 4 places; the RAG and recommendation evaluation in Section 7 remains unmeasured pending system-log data; backend/RAG-scale automated testing was not completed in every development environment (Section 4); the recommendation and itinerary engines are deterministic heuristics, not learned or exactly-optimal models; and no longitudinal evaluation exists.

9. Ethical Considerations

As Stage 1 uses synthetic data, no human participants were involved; in Stage 2, data will be collected from real persons, with contributions restricted to verified users/reviewers. The application never persists a user's precise location beyond the single request that required it and never reads location independently of that request. The system's anti-fabrication grounding discipline (Section 4.1) is intended to reduce, though not eliminate, the risk of the AI presenting an unverified claim as fact, and every stored fact's verification status is exposed rather than hidden. No formal institutional ethics-board approval is claimed at this MCA-project stage; the data-handling practices above are followed as good research practice. No automated plagiarism-detection tool (e.g., Turnitin, Drillbit) has yet been run on this manuscript, which the department's policy (maximum 10% for text-based research) requires before final submission.

Appendix — Questionnaire Overview

The visitor questionnaire designed for Stage 2 data collection from real persons is organized into six sections:

1. Respondent profile — age group, gender, visitor type, travel companions.
2. Destination-visit history — places visited, purpose, first/repeat visit, duration, transport, information sources.

3. Visitor experience — 5-point Likert grid: cleanliness, safety, accessibility, facilities, cultural engagement, overall satisfaction.

4. Destination-selection factors — multi-select: historical/cultural importance, proximity, recommendations, reviews, affordability, natural beauty, and others.

5. Tourism preferences — preferred place types, available time, budget, and importance of personalized recommendation.

6. Visit outcomes — revisit intention, willingness to recommend the place, and open comments.

A seventh, prototype-trial section (rating the usefulness of location-aware discovery, personalized recommendations, the conversational AI assistant and itinerary planning specifically) is included in the design for a future round of data collection once respondents can trial the live system.

10. Conclusion and Future Work

This paper presented an integrated, Pune-focused tourism assistance platform combining a source-verified geospatial knowledge base, a deterministic-first recommendation engine, an itinerary-planning heuristic, and a RAG-grounded conversational AI layer, with eighteen source-verified places and a live-verified geospatial/RAG data layer including one genuinely found-and-fixed indexing defect. Analysis of a 120-response synthetic Stage 1 dataset illustrated strong overall satisfaction and a clear preference for personalized recommendation, alongside crowd comfort as the weakest factor across Pune's heritage sites. Completing a system-log-based RAG/recommendation evaluation (Section 7), running a prototype-trial round of the survey to rate the system's own features directly, expanding the verified dataset and knowledge base beyond their current scope, and closing remaining automated-test gaps are identified as the immediate next steps.

Stage 2 Development (Future Scope). Synthetic data was used in Stage 1 owing to time constraints. In Stage 2 development, the synthetic dataset will be replaced by data collected from real persons, with all reviews and survey responses gathered directly from genuine visitors and the analysis in Section 6 repeated on that data. In addition, Stage 2 will allow only verified users/reviewers to contribute additional information and reviews, both for already-discovered tourist places and for undiscovered but locally popular places, in order to build awareness of lesser-known tourist places in and around Pune.

References

1. Banerjee, A., Banik, P., & Wörndl, W. (2023). A review on individual and multistakeholder fairness in tourism recommender systems. *Frontiers in Big Data*, 6, Article 1168692. <https://doi.org/10.3389/fdata.2023.1168692>
2. Banerjee, A., Satish, A., & Wörndl, W. (2024). *Enhancing tourism recommender systems for sustainable city trips using retrieval-augmented generation* (arXiv:2409.18003). arXiv. <https://doi.org/10.48550/arXiv.2409.18003>
3. Bao, J., Zheng, Y., Wilkie, D., & Mokbel, M. (2015). Recommendations in location-based social networks: A survey. *GeoInformatica*, 19(3), 525–565. <https://doi.org/10.1007/s10707-014-0220-8>
4. Borràs, J., Moreno, A., & Valls, A. (2014). Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, 41(16), 7370–7389. <https://doi.org/10.1016/j.eswa.2014.06.007>



5. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), Article 248. <https://doi.org/10.1145/3571730>
6. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in neural information processing systems* (Vol. 33, pp. 9459–9474). Curran Associates.
7. Orden-Mejía, M., Carvache-Franco, M., Huertas, A., Carvache-Franco, O., & Carvache-Franco, W. (2025). Analysing how AI-powered chatbots influence destination decisions. *PLOS ONE*, 20(3), Article e0319463. <https://doi.org/10.1371/journal.pone.0319463>
8. Sarkar, J. L., Majumder, A., Panigrahi, C. R., Roy, S., & Pati, B. (2023). Tourism recommendation system: A survey and future research directions. *Multimedia Tools and Applications*, 82(6), 8983–9027. <https://doi.org/10.1007/s11042-022-12167-w>
9. Vansteenwegen, P., Souffriau, W., Vanden Berghe, G., & Van Oudheusden, D. (2011). The City Trip Planner: An expert system for tourists. *Expert Systems with Applications*, 38(6), 6540–6546.
10. Zhao, S., King, I., & Lyu, M. R. (2016). *A survey of point-of-interest recommendation in location-based social networks* (arXiv:1607.00647). arXiv. <https://doi.org/10.48550/arXiv.1607.00647>