

TRACE: A Conversational Framework for Sustainable Tourism Recommendation with Agentic Counterfactual Explanations

Ashmi Banerjee

ashmi.banerjee@tum.de

Technical University of Munich
Munich, Germany

Wolfgang Wörndl

woerndl@in.tum.de

Technical University of Munich
Munich, Germany

Adithi Satish

adithi.satish@tum.de

Technical University of Munich
Munich, Germany

Yashar Deldjoo

yashar.deldjoo@poliba.it

Polytechnic University of Bari
Bari, Italy

Abstract

Traditional conversational travel recommender systems primarily optimize for user relevance and convenience, often reinforcing popular, overcrowded destinations and carbon-intensive travel choices. To address this, we present **TRACE** (Tourism Recommendation with Agentic Counterfactual Explanations), a multi-agent, LLM-based framework that promotes sustainable tourism through interactive nudging. TRACE uses a modular orchestrator-worker architecture where specialized agents elicit latent sustainability preferences, construct structured user personas, and generate recommendations that balance relevance with environmental impact. A key innovation lies in its use of *agentic counterfactual explanations* and LLM-driven clarifying questions, which together surface greener alternatives and refine understanding of intent, fostering user reflection without coercion. User studies and semantic alignment analyses demonstrate that TRACE effectively supports sustainable decision-making while preserving recommendation quality and interactive responsiveness. TRACE is implemented on Google's Agent Development Kit, with full code, Docker setup, prompts, and a publicly available demo video to ensure reproducibility. A project summary, including all resources, prompts, and demo access, is available at <https://ashmibanerjee.github.io/trace-chatbot>.

CCS Concepts

• Information systems → Information retrieval; • Computing methodologies → Artificial intelligence.

Keywords

Conversational Recommender Systems, LLMs, Multi-Agent Systems, Sustainable Tourism, Counterfactual Explanations

ACM Reference Format:

Ashmi Banerjee, Adithi Satish, Wolfgang Wörndl, and Yashar Deldjoo. 2026. TRACE: A Conversational Framework for Sustainable Tourism Recommendation with Agentic Counterfactual Explanations. In *Proceedings of the 49th*

International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26), July 20–24, 2026, Melbourne, VIC, Australia. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3805712.3808370>

1 Introduction

Conversational agents are increasingly used to support complex decision-making tasks such as travel planning. Recent advances in LLMs have enabled chat-based travel assistants that can flexibly respond to natural language queries and generate personalized recommendations [19, 31, 33]. However, most existing Conversational Recommender Systems (CRS) for travel primarily optimize for user relevance and convenience, often reinforcing popular, overcrowded destinations and carbon-intensive travel choices. As a result, sustainability considerations, such as environmental impact, congestion, and seasonality, are typically treated as secondary constraints, if considered at all [2, 12, 35].

In this demo paper, we present **TRACE** (*Tourism Recommendation with Agentic Counterfactual Explanations*), an LLM-agent-based conversational recommender system explicitly designed to promote sustainable tourism practices. Unlike traditional chatbots or chat-based travel assistants, TRACE incorporates sustainability objectives from the ground up. The system prioritizes recommendations such as less-visited or emerging destinations over overcrowded hotspots, and favors lower-impact transportation options (e.g., train travel over flights) whenever feasible. Crucially, these sustainability signals are not hard-coded defaults but are inferred through interaction.

The primary goal of this work is *not* to build a state-of-the-art recommender system in terms of predictive accuracy or ranking performance. Instead, we aim to explore the design space and research opportunities enabled by multi-agent LLM architectures for sustainable conversational recommendation. In particular, we focus on how such systems can be structured to *nudge* users toward more sustainable travel choices without violating their explicit preferences or degrading user trust.

TRACE achieves this through a modular, agentic design. Dedicated agents generate targeted clarifying questions to elicit latent sustainability preferences, construct structured user personas, and produce recommendations that jointly optimize for relevance and sustainability signals. A key contribution of the system is its explanation agent, which generates persuasive justifications highlighting the sustainability benefits of recommended options. When users



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

SIGIR '26, Melbourne, VIC, Australia

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2599-9/2026/07

<https://doi.org/10.1145/3805712.3808370>

do not explicitly express sustainability preferences, the system employs counterfactual explanations [20] to gently expose greener alternatives, enabling informed reflection rather than coercive intervention.

We demonstrate TRACE as an interactive system that showcases how agentic LLM pipelines can operationalize responsible nudging in tourism recommendations. By releasing this demo, we aim to stimulate discussion on the role of conversational interfaces, explanations, and counterfactual reasoning in shaping more sustainable user behavior and to encourage further research beyond accuracy-centric evaluation paradigms in recommender systems¹.

Contributions. This paper introduces TRACE, a multi-agent, LLM-based conversational recommender system for sustainable tourism:

- **Modular, multi-agent framework:** Specialized agents handle user modeling, clarifying questions, recommendations, and explanations, integrating digital nudging to promote greener choices.
- **Open-source release and demo:** The full code and framework are publicly available, accompanied by a demo video to facilitate reproducibility and further research².
- **Empirical evaluation:** User studies and semantic metrics show TRACE balances relevance, and sustainability nudging while maintaining interactive response times.

Related Work. Conversational recommender systems have been widely studied for interactive preference elicitation and natural language decision support. Recent LLM-powered, multi-agent approaches leverage specialized agents for user modeling, planning, and explanation, enabling collaborative recommendation and flexible reasoning [22, 25, 27]. However, these systems primarily focus on accuracy and engagement, leaving sustainability and behavior change as secondary concerns [12, 35]. Although recent work [6] has explored multi-agent frameworks to balance relevance and sustainability in tourism recommendations, these approaches do not explicitly incorporate conversational interactions or leverage explanations and counterfactuals to nudge user behavior, and they remain primarily research prototypes with limited production readiness. Explainable and counterfactual recommendation research further explores transparency and bias mitigation, often via minimally contrastive explanations [8, 20, 32, 34, 36]. Relatedly, clarifying questions have been investigated as a way to uncover hidden or incomplete user preferences [9, 28–30], with taxonomies spanning vague, spatial, and temporal dimensions [37] and LLM-based dialogue updates supporting personalization [24]. In parallel, sustainability-aware recommenders apply digital nudging and eco-objectives to guide greener choices [3–5, 21, 26], though most remain single-agent, non-conversational, rely on static sustainability encodings, and often do not leverage LLMs. In contrast, **TRACE** integrates these directions in a multi-agent, LLM-based conversational framework, where agents infer preferences, build structured personas, and explain greener choices as an evolving dialogue objective.

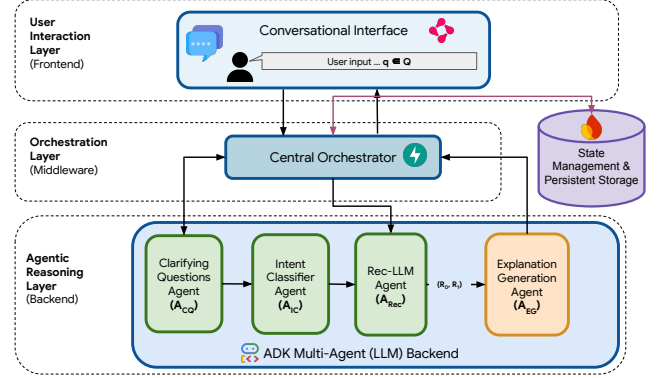


Figure 1: System Architecture of the CRS Chatbot. This diagram illustrates the Orchestrator-Worker paradigm, detailing the flow between the User Interaction layer (Frontend), Orchestration layer (Middleware), and the Agentic Reasoning layer (Backend).

2 TRACE Framework: System Design

The proposed system adopts a modular, microservices-based architecture for delivering sustainable travel recommendations using a Multi-Agent LLM framework. The design follows an Orchestrator-Worker paradigm, where a central middleware component coordinates user interactions and delegates reasoning tasks to specialized agents. The architecture is organized into three conceptual layers:

- **User Interaction Layer (Frontend)**, responsible for conversational interaction and feedback collection.
- **Orchestration Layer (Middleware)**, which manages session state, control flow, and agent execution.
- **Agentic Reasoning Layer (Backend)**, consisting of a sequential pipeline of specialized LLM agents implemented using Google’s Agent Development Kit (ADK) over vertexAI.

Throughout this paper, we define *context* as the information accumulated from a user’s initial query and responses to clarifying questions, which collectively inform the recommendation process by capturing inferred preferences and sustainability considerations. This layered design promotes modularity, scalability, and clear separation of concerns between interaction handling, control logic, and reasoning components. Figure 1 illustrates the overall system architecture, highlighting the interactions between the user interface, orchestrator, and various LLM agents.

2.1 Formalization of the Multi-Agent CRS

We formalize the TRACE Conversational Recommender System (CRS) as a sequential pipeline of transformations over the state space of a user session. Let $q \in Q$ denote the initial natural language query. The system proceeds through the following stages.

2.1.1 Elicitation and Persona Modeling. The Clarifying Question Agent A_{CQ} generates a set of targeted questions $C = \{c_1, c_2, \dots, c_n\}$ to elicit latent user preferences related to sustainable travel:

$$C = A_{CQ}(q). \quad (1)$$

¹Demo video: <https://youtu.be/BdtEFSp42fw>

²<https://ashmibanerjee.github.io/trace-chatbot>

The primary goal of the Clarifying Question Agent is to elicit the user’s sustainability preferences indirectly and their willingness to compromise, for example, by asking whether they would be willing to travel to a lesser-known place rather than a popular (and crowded) destination, thereby improving intent detection. However, if the user’s initial query is vague (“I want to travel in Europe”), then A_{CQ} is also instructed to ask the user about general travel aspects such as their budget and interests. The user’s initial query is provided as input to A_{CQ} , which generates up to 5 clarifying questions, which are then shown to the user one at a time.

Given the user responses Γ to these questions, the Intent Classification Agent A_{IC} maps the interaction into a structured *User Travel Persona* $\mathcal{U}$ and a vector representing the user’s *Willingness to Compromise* (WTC) across sustainability dimensions such as emissions, congestion, and seasonality:

$$(\mathcal{U}, \text{WTC}) = A_{IC}(\Gamma, q). \quad (2)$$

The generated *User Travel Persona* and *Willingness to Compromise* are provided to the recommender to obtain personalized, context-aware recommendations.

2.1.2 Recommendation Generation. The system constructs two candidate recommendation sets. The baseline set R_0 is instructed to prioritize relevance solely to the original query q . In contrast, the sustainable set R_1 incorporates the inferred user persona and sustainability preferences from A_{IC} . Both sets are generated by the Recommender Agent A_{Rec} , a Rec-LLM using few-shot prompting. In particular, R_1 is produced by prioritizing sustainability-related signals S derived from the clarification responses Γ , while preserving alignment with the inferred persona $\mathcal{U}$:

$$R_1 = A_{Rec}(\mathcal{U}, \Gamma, S). \quad (3)$$

2.1.3 Persuasive Explanation and Decision Logic. The Explanation Generation Agent A_{EG} serves as the final decision-making component. It outputs a selected recommendation r^* , a persuasive explanation E , and a counterfactual alternative r_{alt} , explicitly conditioned on the inferred willingness to compromise:

$$(r^*, E, r_{alt}) = A_{EG}(R_0, R_1, \mathcal{U}, \text{WTC}). \quad (4)$$

The agent dynamically adopts one of two rhetorical strategies:

Direct Alignment. If WTC indicates openness to sustainability trade-offs, the agent selects $r^* \in R_1$ and generates an explanation emphasizing the sustainability improvement

$$\Delta S = M(R_1) - M(R_0),$$

where $M(\cdot)$ denotes a vector of sustainability metrics (e.g., CO₂ emissions, visitor pressure, walkability).

Counterfactual Nudging. If WTC indicates resistance to sustainability trade-offs, the agent selects the baseline recommendation $r^* \in R_0$ to preserve user trust, while generating a counterfactual explanation for an alternative $r_{alt} \in R_1$. The explanation is framed conditionally, for example: “Had you expressed interest in lower environmental impact, r_{alt} would have been recommended because ...” This strategy enables implicit nudging toward sustainable options without violating the user’s explicitly stated constraints.

2.2 Implementation

TRACE is implemented as a modular, state-aware pipeline using **Google’s Agent Development Kit (ADK)** [16] on Vertex AI [17], with the gemini-2.5-flash model handling all agentic reasoning [18]. The backend uses an Orchestrator-Worker pattern managed via **FastAPI** [13], with **Google Firestore** [14] maintaining persistent session state (user queries, persona vectors, and WTC). The user interface, built with **Chainlit** [10], supports a seamless Clarifying Question loop and the rendering of recommendations. The stack is containerized with **Docker** and deployed on **Google Cloud Run** [15] for scalable, serverless execution of asynchronous multi-agent workflows.

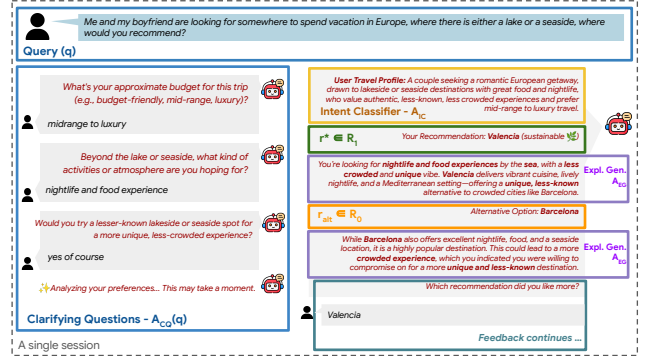


Figure 2: An example TRACE session. The session starts with the user’s initial query, followed by clarifying questions. It concludes with a generated travel profile and a set of recommendations, each accompanied by explanations. The session also collects session-specific feedback, including which recommendations were preferred, how helpful the explanations were, how effective the clarifying questions were, and an option to provide qualitative feedback.

Users can enter free-text queries or select predefined travel scenarios that cover a range of preferences and sustainability considerations (e.g., eco-friendly stays, off-the-beaten-path trips). We use sample queries from SynthTRIPS [7], which provide synthetic travel queries and preferences, to offer users inspiration for their initial inputs.

To prevent misuse, the Clarifying Question Agent (A_{CQ}) enforces guardrails with few-shot instructions, restricting recommendations for single European city trips. Queries outside this scope (for example, “Recommend some movies to watch this weekend”) are flagged as invalid, and users are prompted to stay within city trip recommendations.

2.3 User Journey

In this paper, a *session* denotes a complete interaction between a user and the system, beginning with the user’s initial query and ending with their responses to the feedback questions.

Each session proceeds as follows. Given their query, the user first responds to a set of clarifying questions to construct a travel persona. The system then generates two candidate recommendations:

a baseline recommendation R_0 and a context-aware recommendation R_1 . Based on the user’s willingness to compromise (WTC) on sustainability, the Explanation Generation Agent determines which candidate is presented as the primary recommendation, r^* , and which is shown as the alternative, r_{alt} . When the user indicates no willingness to compromise on sustainability, the agent may assign the baseline recommendation (R_0) as the primary option.

Figure 2 illustrates an example session in which the user expresses a preference for a seaside holiday. In this instance, openness to exploring lesser-known destinations to avoid crowds results in *Valencia* (R_1) being presented as the primary recommendation (r^*), with *Barcelona* (R_0) shown as the alternative (r_{alt}). Had the user instead preferred popular destinations without such flexibility, the roles of the two recommendations would have been reversed.

3 Evaluation and User Insights

We evaluate TRACE using a user study and quantitative alignment analysis to assess its ability to promote sustainable travel choices and support coherent agentic reasoning.

3.1 User Study Design

We recruited 24 participants (58.3% male, 41.7% female) via social media, personal networks, and university mailing lists. Most participants were aged 18–34 (79%), and 54% reported frequent chatbot use. Participants interacted with TRACE using 1–2 travel queries and were encouraged to include challenging or out-of-scope requests (e.g., non-European destinations) to test robustness.

After the interaction, participants rated the system on a 5-point Likert scale (1: Not at all, 5: Extremely) [23] across three dimensions: *quality of clarifying questions*, *persuasiveness of explanations*, and *extent of choice reconsideration*. After filtering incomplete and out-of-scope responses, we analyzed $N = 107$ valid conversations.

The evaluation examines whether TRACE can nudge users toward sustainable choices without being coercive, while maintaining alignment across agents and with user intent. We formalize these objectives through the following research questions (RQs):

- (1) **User Feedback and Interaction Analysis (RQ1):** *How effectively does the system promote sustainable choices while preserving user trust and engagement?*
- (2) **System Alignment and Semantic Metrics (RQ2):** *To what extent does the system maintain alignment across its internal agents and with user intent?*
- (3) **System Latency (RQ3):** *Does TRACE achieve response times suitable for real-time interaction despite the computational overhead of multiple agents?*

3.2 RQ1: User Feedback and Interaction Analysis

RQ1 investigates whether TRACE promotes sustainable choices while preserving user trust and engagement. Results from the user study (Figure 3) show high acceptance of the system’s conversational interaction. When presented with a primary and alternative recommendation, 79.1% of users selected the primary option (r^*), which in 75.5% of sessions corresponded to the context-aware (typically more sustainable, i.e., R_1) recommendation. In sessions where the baseline (R_0) was presented as the primary option (r^*), 16.7%

of users selected the context-aware alternative (R_1), indicating a modest nudging effect.

User feedback further supports the effectiveness of the interaction design: 55.2% of users rated the clarifying questions and 65.4% rated the explanations as *Very Well* or *Extremely Well*. Moreover, around 60% of users reported some degree of reconsideration of their initial choice, suggesting that TRACE encourages reflection while preserving user agency. Overall, these findings indicate that TRACE can support sustainable recommendations without compromising user engagement or trust.

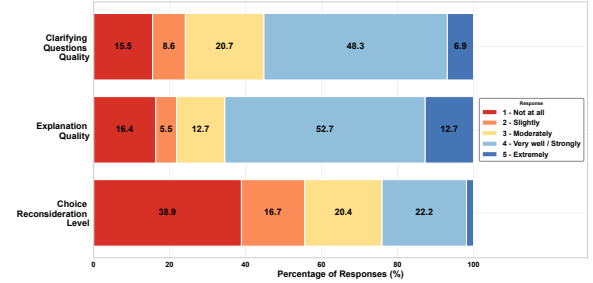


Figure 3: Combined feedback distribution across rating categories for Clarifying Question Quality, Explanation Quality, and Choice Reconsideration Level.

3.3 RQ2: System Alignment and Semantic Metrics

RQ2 examines whether TRACE maintains alignment between the conversation context, the structured user model, and the explanations generated by the agentic pipeline. Using semantic similarity metrics computed with the all-MiniLM-L6-v2 model, we measure coherence across these components. The explanations generated by A_{EG} show high similarity to the full conversation (User Query q + Clarifying Questions), with a mean score of 0.7033, indicating strong contextual grounding. The output of the Intent Classifier (User Persona, Travel Intent, and WTC) also aligns closely with the conversation (0.7883), confirming accurate intent capture. Finally, the high similarity between the Intent Classifier output and the generated explanations (0.7437) demonstrates that A_{EG} consistently conditions its reasoning on the inferred user model and sustainability dimensions. Overall, these results indicate strong alignment across TRACE’s internal agents and with user intent.

3.4 RQ3: System Latency

We assess whether TRACE can operate within latency bounds suitable for interactive use despite its multi-agent design. Following the clarifying question phase, TRACE requires an average of 23 seconds to generate user profiles, recommendations, and explanations, with a maximum latency of 38 seconds. These results show that TRACE’s modular architecture does not introduce prohibitive overhead, supporting the practical feasibility of multi-agent conversational recommendation. Moreover, its latency is competitive with that of contemporary multi-agent frameworks [11], which often exceed 40 seconds. Overall, TRACE effectively balances the

computational cost of agentic reasoning with the need for accurate and personalized tourism recommendations.

4 Conclusion

We presented **TRACE**, a multi-agent LLM-based conversational recommender system for sustainable tourism, combining user modeling, clarifying questions, recommendations, and explanations to nudge users toward greener choices while preserving user trust. User studies and semantic metrics show TRACE balances relevance, trust, and sustainability, nudging with interactive response times.

Due to a Chainlit vulnerability discovered on 21 January, 2026 [1], the live app is currently offline, but the code, Docker setup, prompts, and a demo video are publicly available for reproducibility. The system ran successfully from January 16–31, 2026, and can be set up locally using our resources.

Currently limited to single European city travel, TRACE will be extended in future work to multi-city, multi-day itineraries. Overall, it demonstrates the feasibility of combining multi-agent reasoning, LLMs, and sustainability-aware conversational recommendation in a practical, reproducible framework. At the same time, TRACE raises an important *sustainability paradox*. By consistently nudging users toward less popular “hidden gems,” the system may inadvertently create new hotspots, shifting rather than reducing overtourism. Addressing this requires dynamic, adaptive recommendation strategies that account not only for individual preferences but also for destination-level impacts. Future work should therefore incorporate real-time signals, such as destination capacity, environmental indicators, and user feedback, to continuously refine recommendations and mitigate such unintended consequences.

GenAI Usage Disclosure

We used ChatGPT, Claude, and Gemini for code suggestions, and Grammarly for language refinement; all outputs were critically reviewed, and we take full responsibility for the final version.

Acknowledgments

We thank the Google AI/ML Developer Programs team for supporting us with Google Cloud Credits.

References

- [1] Ionut Arghire. 2026. Chainlit Vulnerabilities May Leak Sensitive Information. <https://www.securityweek.com/chainlit-vulnerabilities-may-leak-sensitive-information/>. Accessed 2026-02.
- [2] Ashmi Banerjee. 2023. Fairness and sustainability in multistakeholder tourism recommender systems. In *Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*. 274–279.
- [3] Ashmi Banerjee, Melih Mert Aksoy, and Wolfgang Wörndl. 2025. SmartSustain Recommender System: Navigating Sustainability Trade-offs in Personalized City Trip Planning. *arXiv preprint arXiv:2510.17355* (2025).
- [4] Ashmi Banerjee, Tunar Mahmudov, Emil Adler, Fitri Nur Aisyah, and Wolfgang Wörndl. 2025. Modeling sustainable city trips: integrating CO₂e emissions, popularity, and seasonality into tourism recommender systems. *Information Technology & Tourism* 27, 1 (2025), 189–226.
- [5] Ashmi Banerjee, Tunar Mahmudov, and Wolfgang Wörndl. 2024. Green Destination Recommender: A Web Application to Encourage Responsible City Trip Recommendations. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*. 486–490.
- [6] Ashmi Banerjee, Adithi Satish, Fitri Nur Aisyah, Wolfgang Wörndl, and Yashar Deldjoo. 2025. Collab-REC: An LLM-based Agentic Framework for Balancing Recommendations in Tourism. *arXiv preprint arXiv:2508.15030* (2025).
- [7] Ashmi Banerjee, Adithi Satish, Fitri Nur Aisyah, Wolfgang Wörndl, and Yashar Deldjoo. 2025. SynthTRIPS: A Knowledge-Grounded Framework for Benchmark Data Generation for Personalized Tourism Recommenders. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 3743–3752.
- [8] Oren Barkan, Veronika Bogina, Liya Gurevitch, Yuval Asher, and Noam Koenigstein. 2024. A counterfactual framework for learning and evaluating explanations for recommender systems. In *Proceedings of the ACM Web Conference 2024*. 3723–3733.
- [9] Keping Bi, Qingyao Ai, and W Bruce Croft. 2021. Asking clarifying questions based on negative feedback in conversational search. In *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*. 157–166.
- [10] Chainlit. 2026. Chainlit: Get Started – Overview. <https://docs.chainlit.io/get-started/overview>.
- [11] Philip Drammeh. 2025. Multi-Agent LLM Orchestration Achieves Deterministic, High-Quality Decision Support for Incident Response. *arXiv preprint arXiv:2511.15755* (2025).
- [12] Jiabao Fang, Shen Gao, Pengjie Ren, Xiuying Chen, Suzan Verberne, and Zhaochun Ren. 2024. A multi-agent conversational recommender system. *arXiv preprint arXiv:2402.01135*.
- [13] FastAPI. 2026. FastAPI Documentation. <https://fastapi.tiangolo.com/>.
- [14] Firebase and Google Cloud. 2026. Cloud Firestore Documentation. <https://firebase.google.com/docs/firestore>.
- [15] Google Cloud. 2026. Cloud Run Documentation. <https://cloud.google.com/run>.
- [16] Google Cloud. 2026. Overview of Agent Development Kit. <https://docs.cloud.google.com/agent-builder/agent-development-kit/overview>.
- [17] Google Cloud. 2026. Vertex AI Platform. <https://cloud.google.com/vertex-ai>.
- [18] Google DeepMind. 2024. *Gemini 2.5: Technical Report*. Technical Report. Google DeepMind. https://storage.googleapis.com/deepmind-media/gemini/gemini_v2_5_report.pdf.
- [19] Shengyu Gu. 2024. A survey of large language models in tourism (Tourism LLMs). *Preprint on Qeios* (2024).
- [20] Riccardo Guidotti. 2024. Counterfactual explanations and how to find them: literature review and benchmarking. *Data Mining and Knowledge Discovery* 38, 5 (2024), 2770–2824.
- [21] Haya Halimeh and Oliver Müller. 2025. Towards Greener Choices: Decision Information Nudging for Sustainability-Aware Recommender Explanations. In *International Workshop on Recommender Systems for Sustainability and Social Good*. Springer, 27–42.
- [22] Chengkai Huang, Junda Wu, Yu Xia, Zixu Yu, Ruhan Wang, Tong Yu, Ruiyi Zhang, Ryan A Rossi, Branislav Kveton, Dongruo Zhou, et al. 2025. Towards agentic recommender systems in the era of multimodal large language models. *arXiv preprint arXiv:2503.16734* (2025).
- [23] Ankur Joshi, Saket Kale, Satish Chandel, and D Kumar Pal. 2015. Likert scale: Explored and explained. *British journal of applied science & technology* 7, 4 (2015), 396–403.
- [24] Sara Kemper, Justin Cui, Kai Dicarantonio, Kathy Lin, Danjie Tang, Anton Korikov, and Scott Sanner. 2024. Retrieval-augmented conversational recommendation with prompt-based semi-structured natural language state tracking. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2786–2790.
- [25] Reza Yousefi Maragheh and Yashar Deldjoo. 2025. The Future is Agentic: Definitions, Perspectives, and Open Challenges of Multi-Agent Recommender Systems. *arXiv preprint arXiv:2507.02097* (2025).
- [26] Noemi Mauro, Livio Scarpinati, Fabio Ferrero, Angelo Geninatti Cossatin, and Claudio Mattutino. 2024. Point-of-Interest Recommender Systems: Nudging towards Sustainable Tourism. In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*. 491–495.
- [27] Qiyao Peng, Hongtao Liu, Hua Huang, Qing Yang, and Minglai Shao. 2025. A survey on llm-powered agents for recommender systems. *arXiv preprint arXiv:2502.10050* (2025).
- [28] Xuhui Ren, Hongzhi Yin, Tong Chen, Hao Wang, Zi Huang, and Kai Zheng. 2021. Learning to ask appropriate questions in conversational recommendation. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*. 808–817.
- [29] Ivan Sekulić, Mohammad Aliannejadi, and Fabio Crestani. 2021. Towards Facet-Driven Generation of Clarifying Questions for Conversational Search (ICTIR '21). Association for Computing Machinery, New York, NY, USA, 167–175. doi:10.1145/3471158.3472257
- [30] Ivan Sekulić, Weronika Łajewska, Krisztian Balog, and Fabio Crestani. 2024. Estimating the usefulness of clarifying questions and answers for conversational search. In *European Conference on Information Retrieval*. Springer, 384–392.
- [31] Zijian Shao, Jiancan Wu, Weijian Chen, and Xiang Wang. 2025. Personal Travel Solver: A Preference-Driven LLM-Solver System for Travel Planning. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 27622–27642.
- [32] Juntao Tan, Shuyuan Xu, Yingqiang Ge, Yunqi Li, Xu Chen, and Yongfeng Zhang. 2021. Counterfactual explainable recommendation. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 1784–1793.

- [33] Ke Wang, Shuai Yan, Haoran Yuan, Yanling Huang, Yuhang Wu, Fei Li, Shengying Yang, and Huan Deng. 2025. Toward Interpretable and Persistent Personalization: A Memory-Augmented Agent Framework for LLM-Based Travel Planning. *IEEE Access* 13 (2025), 193125–193141.
- [34] Xiangmeng Wang, Qian Li, Dianer Yu, Qing Li, and Guandong Xu. 2024. Counterfactual explanation for fairness in recommendation. *ACM Transactions on Information Systems* 42, 4 (2024), 1–30.
- [35] Zhefan Wang, Yuanqing Yu, Wendi Zheng, Weizhi Ma, and Min Zhang. 2024. Macrec: A multi-agent collaboration framework for recommendation. (2024), 2760–2764.
- [36] Dianer Yu, Qian Li, Xiangmeng Wang, Qing Li, and Guandong Xu. 2023. Counterfactual explainable conversational recommendation. *IEEE Transactions on Knowledge and Data Engineering* 36, 6 (2023), 2388–2400.
- [37] Hamed Zamani, Susan Dumais, Nick Craswell, Paul Bennett, and Gord Lueck. 2020. Generating clarifying questions for information retrieval. In *Proceedings of the web conference 2020*. 418–428.