

Building Responsible Multi-Agent Recommender Systems with Gemini & Google ADK



Ashmi Banerjee

PhD Candidate @ TU Munich

AI / Cloud -GDE

@ashmi_banerjee

About Me



Photo from IWD Munich 2023



PhD Candidate at TU Munich (currently transitioning)

Researching Tourism Recommender Systems
MSc. TU Munich, Germany (2019)



Professional Experience

Previously: Data Scientist @ Deutsche Telekom, Intern @ MPI-SWS (2018-19)



AI/Cloud Google Developer Expert (GDE)

WTM Ambassador



Loves traveling (**45+ countries; 5 continents**) & outdoor activities



Today's Agenda

PART 01



Evolution

RecSys Architectural Shift

1. Traditional RecSys
2. LLM-Powered RS
3. **Agentic Recommendation**

PART 02



Research Systems

Frameworks & Implementations

Collab-REC

Collaborative agent-based recommendation model

TRACE

Transparent & accountable recommendation trajectory

PART 03

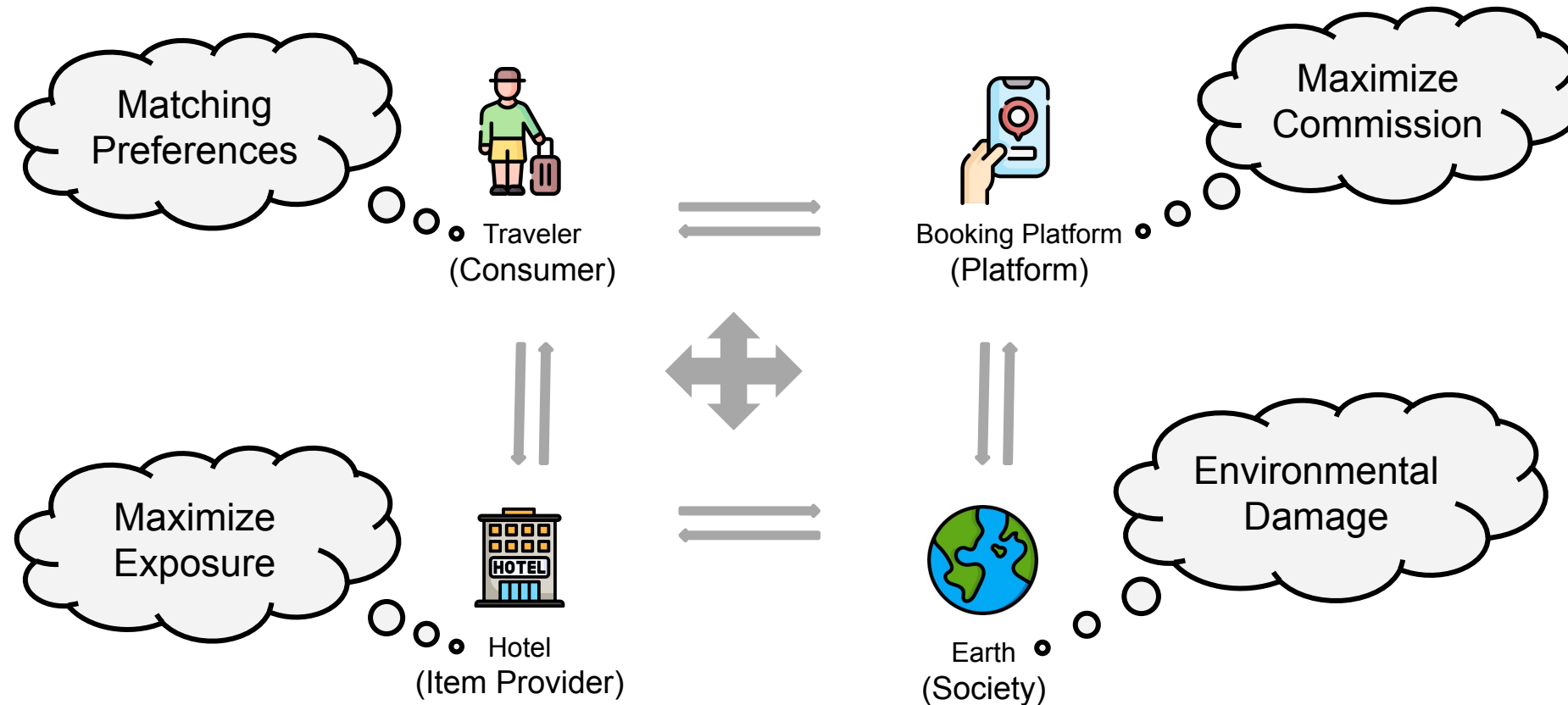


Challenges

Agentic Systems & Responsibility

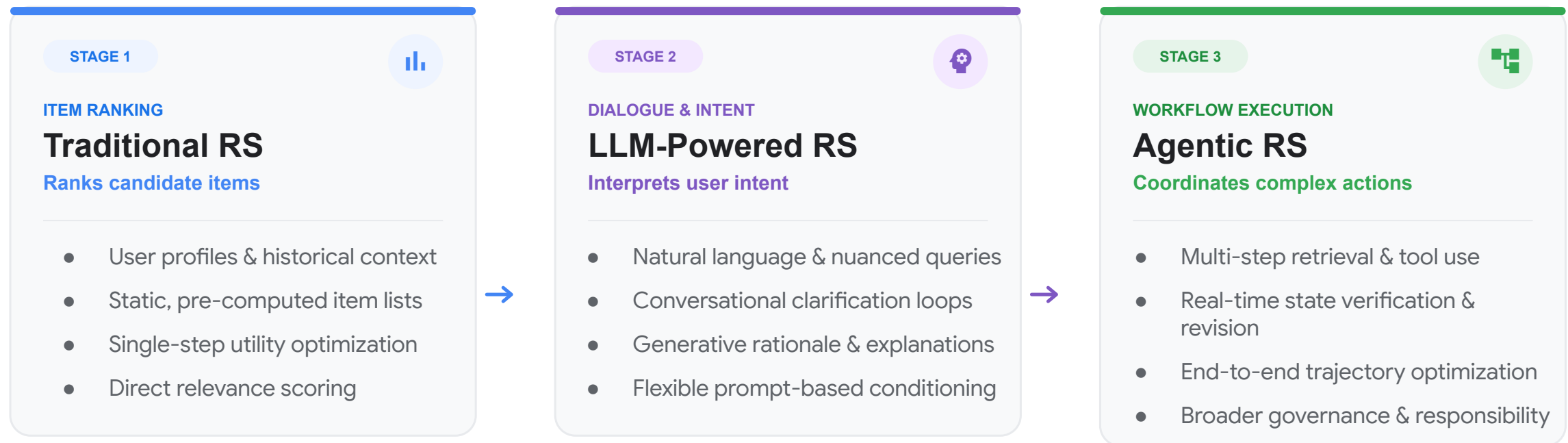
- **Multi-Stakeholder Dynamics:** Balancing traveler, provider & platform needs
- **Governance & Trust:** Ensuring responsible, safe agentic behaviors

A Recommender System: Multi-Stakeholder Scenario



From Traditional RS to Agentic Recommendation

The structural evolution from static matrix ranking to interactive intent parsing and dynamic action execution.



LLMs change the interface.

Agents change the decision process.



“Plan a quiet, walkable weekend under €500.”

LLM-POWERED (monolithic)

Understands nuance
Asks clarifying questions
Explains in natural language

AGENTIC

Retrieve

Verify

Coordinate

Revise

**Evaluate the trajectory,
not only the final answer.**

The Case for Agentic, Moderator-Mediated Design

BASED ON THE WORK

Collab-REC: An LLM-based Agentic Framework for Balancing Recommendations in Tourism

Ashmi Banerjee¹, Adithi Satish¹, Fitri Nur Aisyah¹, Wolfgang Wörndl¹, Yashar Deldjoo²

Accepted at ACM Transactions on Recommender Systems (TORS) 2026

¹ Technical University of Munich, Germany ² Polytechnic University of Bari, Italy



Scan to read paper
<https://bit.ly/collab-rec-paper>

Why Monolithic LLMs Are Ill-Suited

Monolithic LLM Pipeline

Single Prompt: personalize + rank + sustainability + popularity

Opaque ranking → popularity dominates

- **Objective collapse:** goals compressed into one prompt → popularity wins
- **Hallucination:** fabricated destinations or sustainability metrics
- **No auditability:** trade-offs hidden deep inside the black-box model



Objective Collapse

Competing goals compressed into one prompt → model defaults to statistically dominant pattern: globally popular cities



Popularity Bias

Even when explicitly prompted for sustainable alternatives, LLMs frequently revert to Paris, London, Amsterdam



No Auditability

End-to-end generation obscures intermediate reasoning — impossible to explain or verify trade-offs to stakeholders

The Case for Agentic, Moderator-Mediated Design

Distribute objectives across specialized agents — aggregate via a transparent, deterministic moderator



Objective Isolation

Each agent optimizes one stakeholder goal in full. A Popularity Agent deliberately searches the long tail; a Personalization Agent enforces hard filters.



Transparent Negotiation

A non-LLM moderator audits, penalizes, and rewards each agent's behavior — surfacing implicit trade-offs rather than hiding them.



Bias Mitigation by Design

Three agents pulling in different directions. Their iterative compromise, orchestrated by the moderator, naturally dampens the popularity bias.

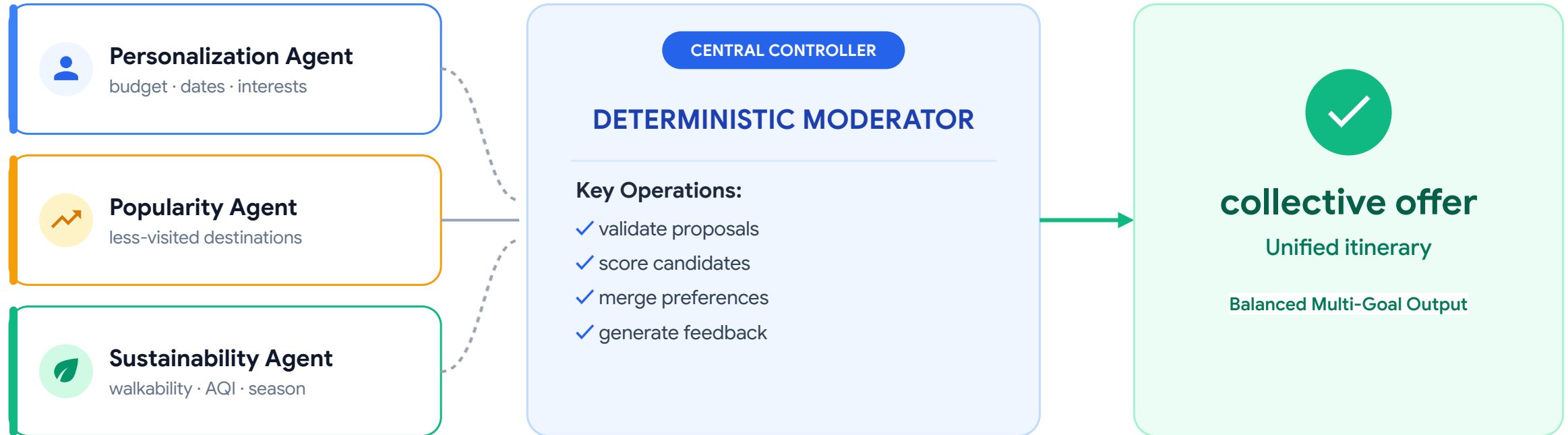


Graceful Extensibility

New stakeholder roles (safety, accessibility) can be added as additional agents without retraining — echoing calls for modular Gen-RecSys.

Collab-REC: Framework Overview

A Google ADK-Based Multi-Agent Tourism Recommender for Mitigating Popularity Bias in European City Trips



Insight: Coordination occurs through shared moderator state — not direct agent debate.

Empirical Evidence from Collab-REC

MAIN RESULT

Multi-round coordination improves grounded relevance against a single-agent baseline.

Shifts exposure significantly toward the mid/long tail.

900

queries

6

LLM families

3–5

rounds to stabilize

81.5%

Catalog Coverage

Claude (vs 66% single-shot)

**↑ Long-tail
Diversity**

Shifts away from Paris / London / Amsterdam

~87s

Practical Latency

With patience early-stop at round 4

0 halluc.

After Grounding

Moderator eliminates out-of-catalog cities

Can an explanation change the choice without taking the choice away?

BASED ON THE WORK

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

Ashmi Banerjee¹, Adithi Satish¹, Wolfgang Wörndl¹, Yashar Deldjoo²

 ACM SIGIR 2026
Melbourne, Australia

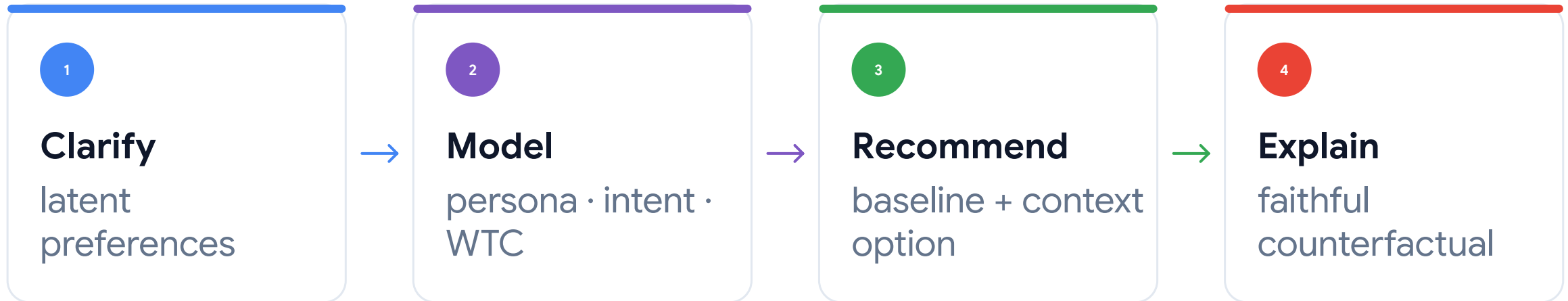
¹ Technical University of Munich, Germany ² Polytechnic University of Bari, Italy



Scan to read paper
<https://bit.ly/trace-acm-paper>

TRACE connects elicitation, user modeling, recommendation, and explanation

Built with Gemini and Google ADK



GOOGLE ADK

orchestrator-worker composition · session state · Gemini agents · reproducible prompts

The explanation is conditioned on the conversation and structured user model.

A counterfactual makes the trade-off inspectable

USER

“A lively seaside city in May — but I dislike crowds.”

1

CLARIFY

Would you accept a lesser-known destination?

2

INFER

Openness to lower crowding: high

3

RECOMMEND

Valencia as primary · Barcelona as alternative

4

COUNTERFACTUAL

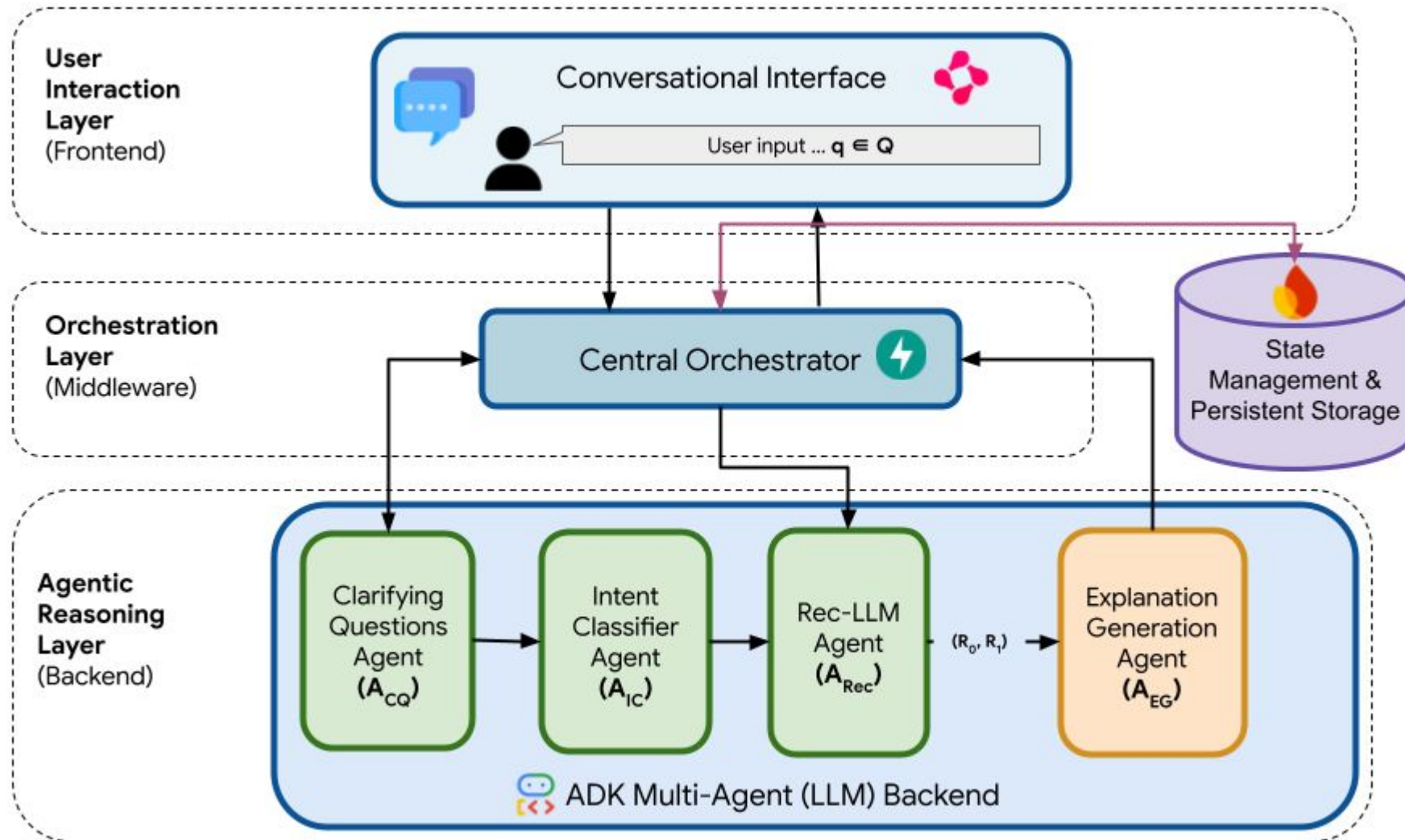
If popularity mattered more, Barcelona would rank first.

5

REVISE

User can reject the inference and rerun the decision.

TRACE: Framework Overview



Main Results: TRACE nudges reflection while preserving a visible alternative

107

valid conversations

75.5%

context-aware option primary

~60%

some reconsideration

23s

average latency

Semantic alignment

conversation ↔ user model

**0.7883**

user model ↔ explanation

**0.7437**

conversation ↔ explanation

**0.7033**

Interpret carefully

The results show coherent conditioning and user reflection — not proof of long-term behavior change or net environmental benefit.

Open Challenges at the Intersection of LLMs, Multi-Stakeholder RecSys & Responsible AI

Evaluation Benchmarks

No standardized benchmarks that jointly capture user-centric AND societal metrics (overtourism, equity of exposure, environmental impact).

Adaptive Weighting

Future work: learned weighting that responds to query context, platform policy, or real-time destination capacity signals.

Human-in-the-Loop

Human oversight and governance mechanisms for agentic recommenders remain an open problem.

Agent Stability & Stochasticity

Stable behavior across rounds is hard to guarantee with stochastic LLMs. Moderator feedback is critical but not a complete solution.

Beyond Tourism

The argument generalizes to domains where recommendations have societal externalities: healthcare, mobility, and cultural heritage.

Other Challenges of Agentic Design

Not just limited to Recommender Systems



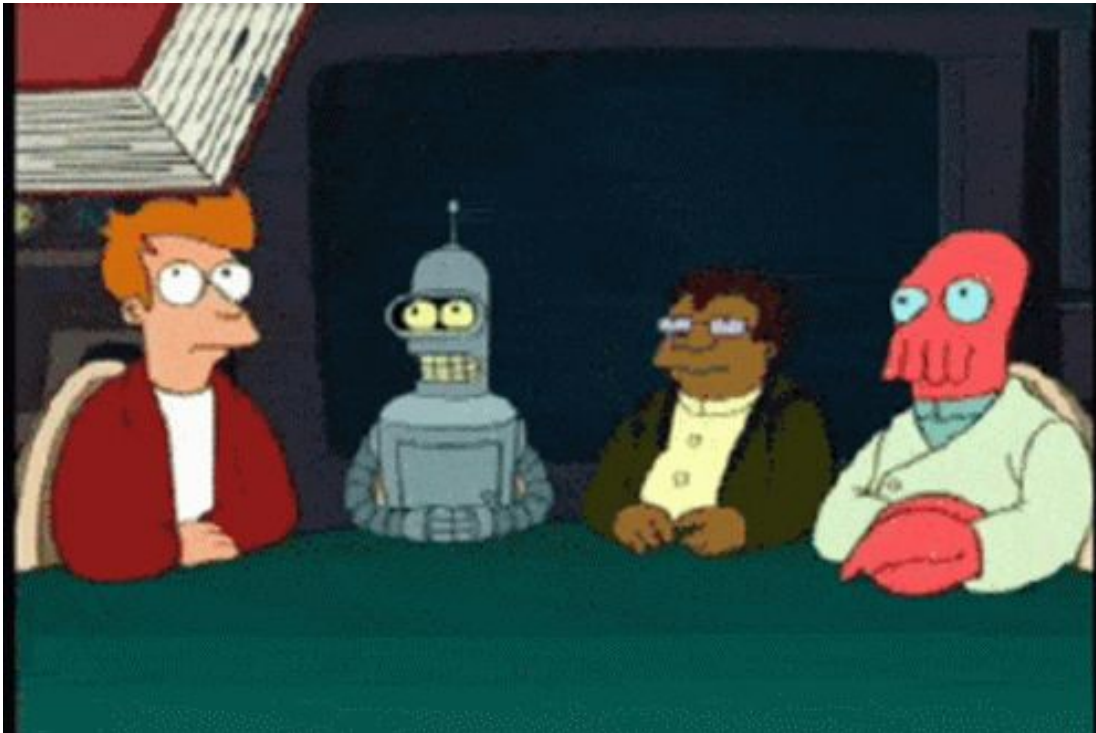
OPENAI AGENT HACKED INTO AUSTRALIA'S NATIONAL HEALTHCARE SYSTEM



CNN



Recap



01. Monolithic LLMs are structurally mismatched

Opaque generation, objective collapse, and popularity bias are architectural problems in tourism context.

02. Agentic, moderator-mediated design

Specialized agents with transparent aggregation support grounded, auditable, sustainability-aware recommendations.

03. Collab-REC/ TRACE

The Proof of Concept

Thank You! Any Questions?



Ashmi Banerjee

PhD Candidate @ TU Munich

AI / Cloud GDE
WTM Ambassador

@ashmi_banerjee



Scan for Talk Feedback

Please share your thoughts and suggestions!

bit.ly/ashmib-feedback