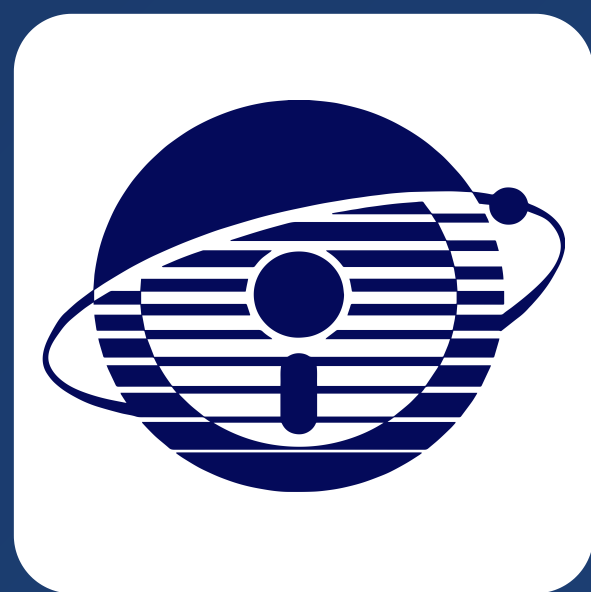




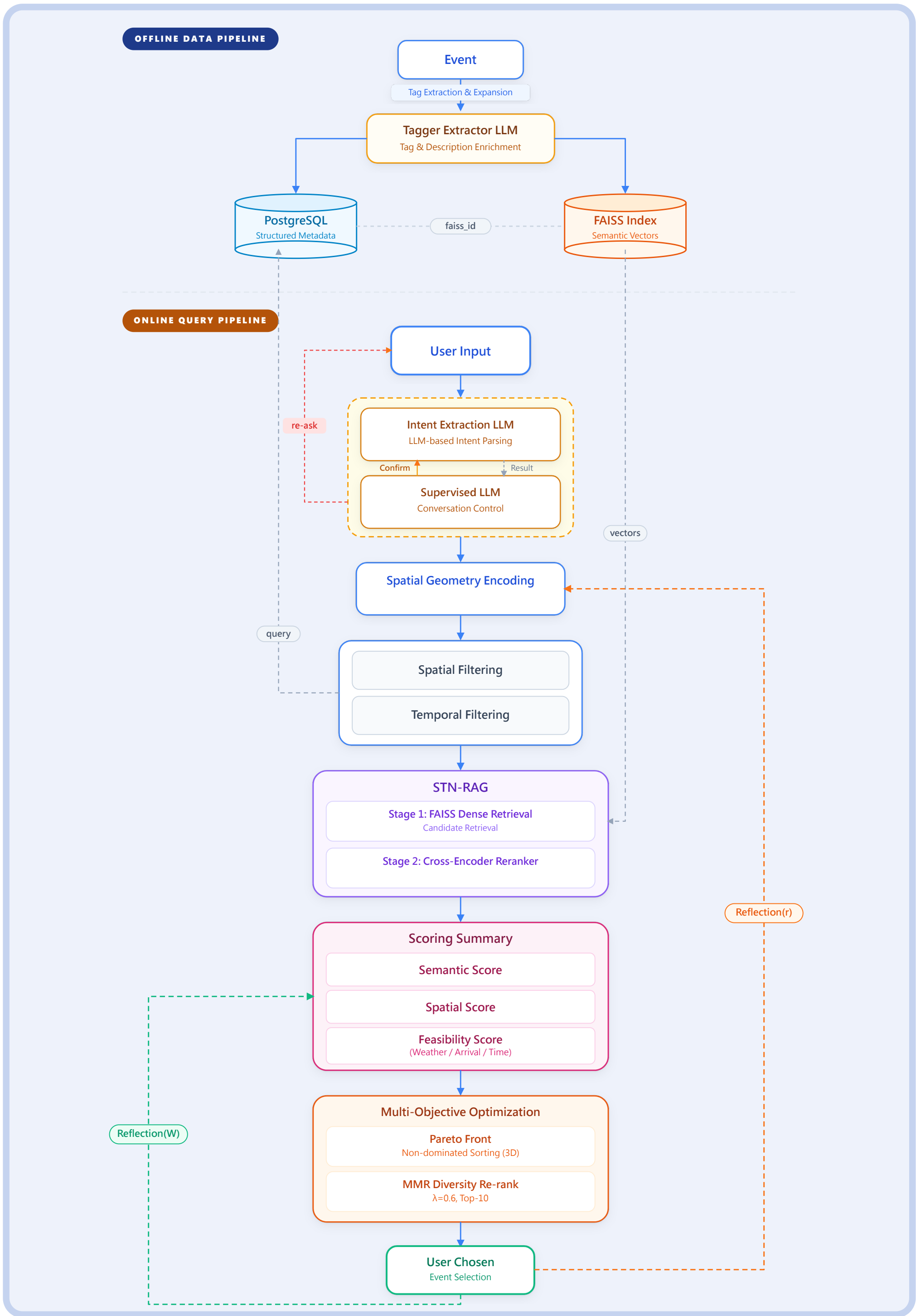
EventPilot: Natural Language Geo-Event Search with LLM-RAG

Group C-8 | Advisor: Prof. Cheng-Te Li | Member: Hao-Cheng Wang

1 MOTIVATION

Urban event data is highly fragmented, and conventional maps fail to interpret contextual needs like "indoor on rainy days." Meanwhile, general LLMs lack spatial grounding, leading to navigation hallucinations. EventPilot bridges semantic understanding and spatial computation, translating vague natural language requests into computable geo-queries and multi-objective ranking to deliver precise, context-aware event recommendations.

2 SYSTEM ARCHITECTURE



3 CORE TECHNOLOGIES

STN-RAG

Converts vague spatial terms like "nearby" into point/line/polygon buffer queries. Combines two-stage retrieval (FAISS recall + cross-encoder reranking) with weather/traffic constraints and MMR diversity

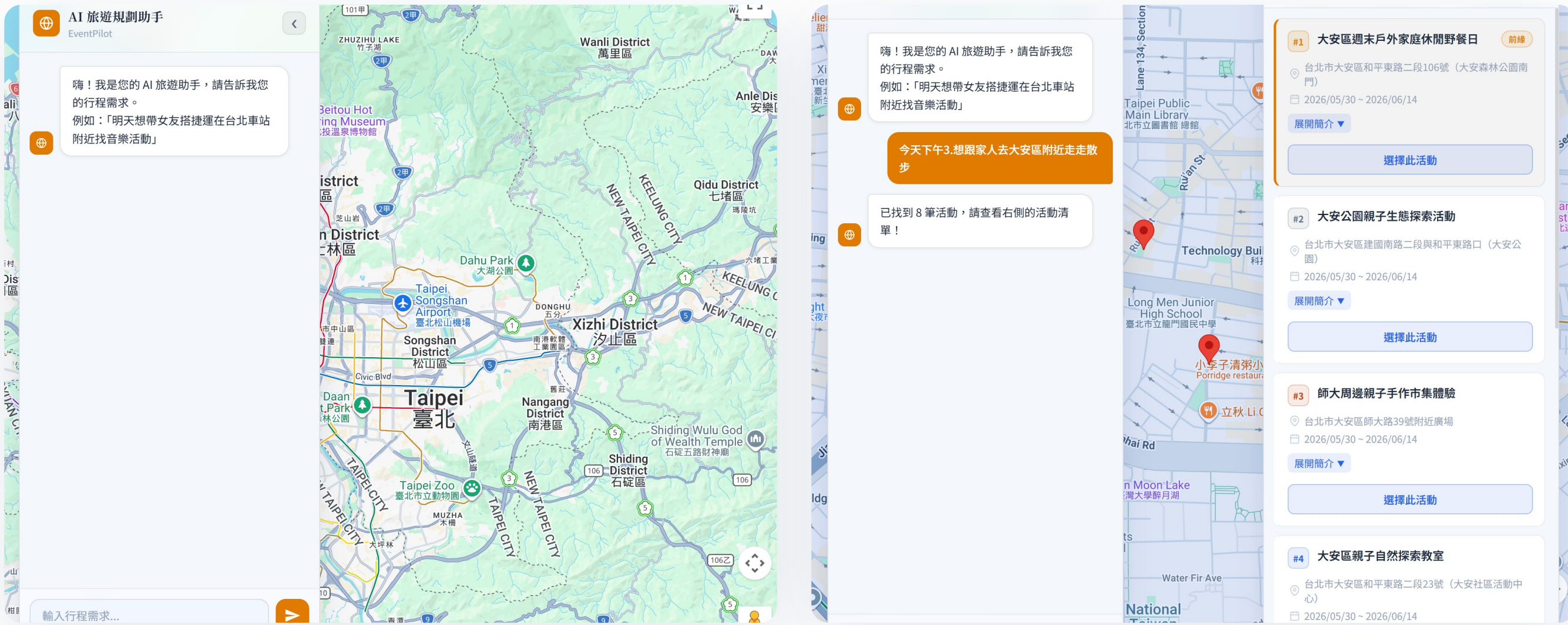
Multi-Objective Optimization

Uses non-dominated sorting to find the Pareto front across three goals — semantic relevance, spatial efficiency, and feasibility — so no single metric dominates

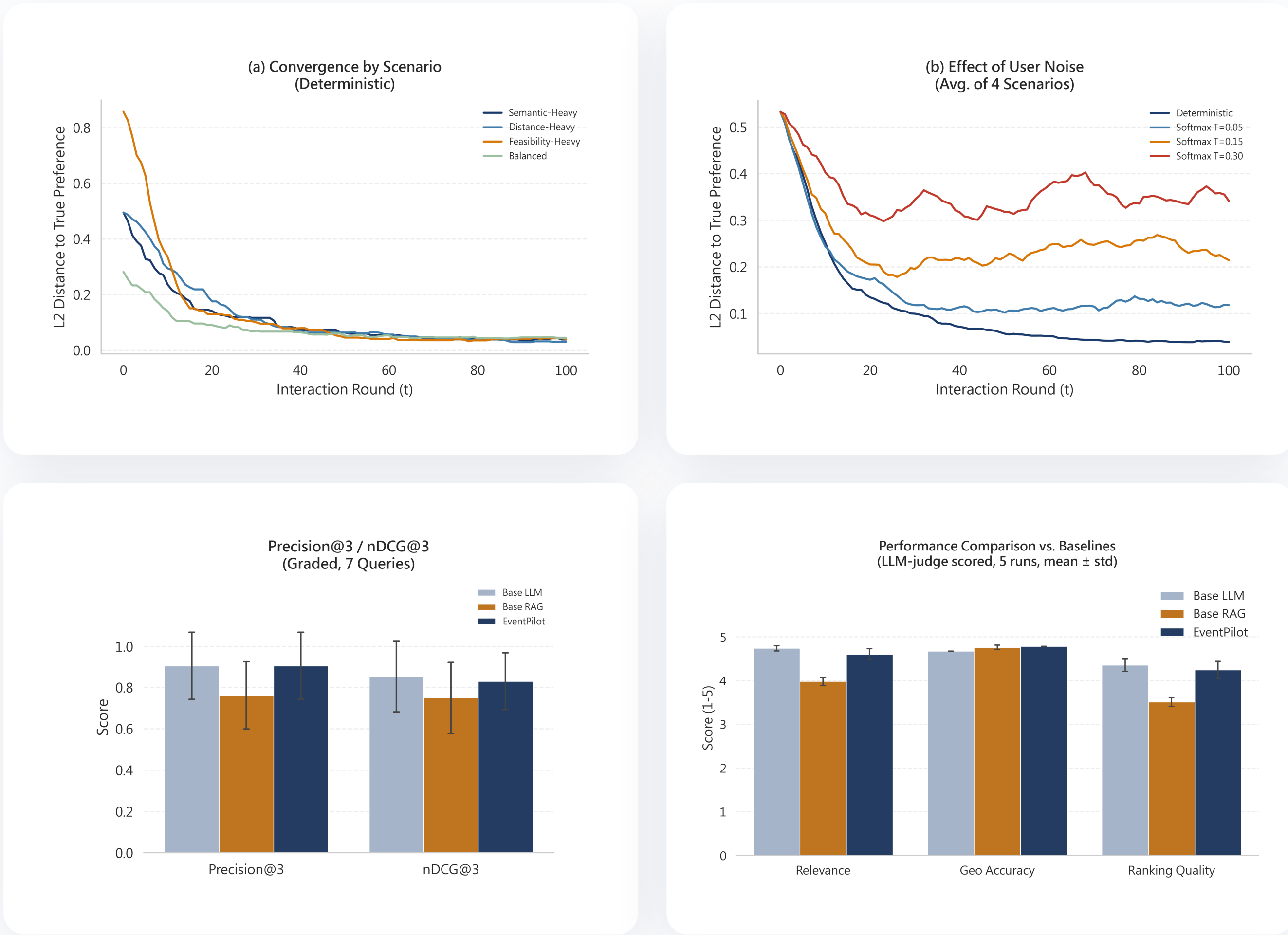
Reflective Agent

Adjusts ranking weights W , spatial tolerance r , and interest vector based on user accept/reject behavior, converging toward personal preferences

4 SYSTEM DEMO



5 QUANTITATIVE EVALUATION



Ablation	P@3	$\Delta P@3$	nDCG@3	$\Delta nDCG$
Full Model	0.905	—	0.830	—
– Cross-Encoder Reranker	0.857	+0.048	0.814	+0.016
– Weather × Outdoor Coef	0.810	+0.095	0.755	+0.075
– Pareto Front	0.857	+0.048	0.766	+0.064
– Arrival Score	0.857	+0.048	0.783	+0.047

6 CONCLUSION & FUTURE WORK

EventPilot uses an STN-RAG architecture to translate spoken queries into point/line/polygon geo-buffer queries, combining two-stage retrieval, a weather × venue interaction factor, Pareto ranking, and a Reflective Agent for personalized weight tuning. Its weights converge stably across scenarios, even under user noise. On a 7-query graded evaluation, the system reaches **P@3 = 0.905**, **nDCG@3 = 0.830**, a 10.7-point nDCG gain over Base RAG, outperforming the baseline overall.

Short-Term

Integrate real-world data and users for live deployment

Mid-Term

Replace remaining LLM steps with lightweight fine-tuned models to cut cost

Long-Term

Expand coverage island-wide with a long-term scalable architecture

