Project Summary

Conversational Package Queries for Subjective Summarization
Fall 2025 / Spring 2026 Honors Thesis – Aryan Deshpande, Vasilis Vittis

Overview

Modern users increasingly seek *personalized* text summaries based on their *subjective intents*, refined interactively over multi-turn dialogue [1]. For example:

Turn 1: "Summarize the art culture of US state X." $\rightarrow$ Summary$_1$ (...)
Turn 2: "Can you refine the summary by placing greater emphasis on fine arts?" $\rightarrow$ Summary$_2$ (...)

Current *Conversational Recommender Systems (CRS)* powered by LLM agents have made rapid progress toward this goal, yet recent surveys and user studies still report challenges in *subjectivity alignment* (i.e., matching each user's personal preferences), *faithful control* (i.e. avoiding hallucinations), and overall user satisfaction [2–4]. At the same time, *package queries* [5] give the kind of formal, verifiable framework that current LLM-style systems lack: they provide explicit numeric constraints and an objective function, so the resulting set (*package*) is provably within bounds instead of being a best-effort guess.

Existing work, SuDocu [6], shows how the package query approach can be applied to text summarization: The user first clicks a handful of sentences to form an *example summary* that reflects their intent (e.g., "Summarize the art culture of US state X"). The summarization task is framed as a *package query*: each candidate sentence is an item, the *summary is a package* and the content properties of the example summary—such as how often it mentions key topics—are converted into numeric bounds that constrain the query. The idea of learning-by-example is suitable when the *intent* remains the same across different sources; in that case, the *same package query* can be executed on a new document (e.g., "Summarize the art culture of US state Y") without gathering additional examples. Later, SUBSUME [7] showed that *example-driven* methods beat query-only baselines on subjective intents; among the example-driven systems, SBERT-EX ranks first and SuDocu second, leaving plenty of room for improvement in subjective summarization and illustrating the importance of further research in this area [7].

The proposed research aims to extend prior work over *CRS* settings. SuDocu shows that package queries can create a summary from a single user request, but it does not support follow-up refinements. Each time a user refines their intent (e.g., "...greater emphasis on fine arts?"), they must create a new example summary from scratch. Although this strategy captures user intent effectively, it is cumbersome in multi-turn dialogue, where users prefer brief edits and expect instant updates. *Our goal is to extend* SuDocu *to support conversational, multi-turn intent refinement, enabling interactive, exploratory "what-if" summaries without requiring new example clicks.*

Intellectual Merit

The project extends in-database prescriptive analytics to dynamic, conversational settings by introducing (i) a formal approach for feedback-driven bound adaptation; (ii) incremental package maintenance for successive conversational turns; and (iii) comprehensive evaluations on multi-turn subjective summarization.

1. **Intent-to-Constraint Mapping (I2C).** We will translate brief intent refinements into small, auditable updates of package-query bounds and objective weights; explain changes, and flag infeasibility.
2. **Incremental Package Maintenance.** We will contribute to warm-started, progressive optimization that reuses the previous solution, delivering fast updates with improvement guarantees and full provenance.

References

[1] H. Zhang, X. Liu, and J. Zhang, "Summit: Iterative text summarization via chatgpt," 2023. [Online]. Available: https://arxiv.org/abs/2305.14835

[2] L. Wu, Z. Zheng, Z. Qiu, H. Wang, H. Gu, T. Shen, C. Qin, C. Zhu, H. Zhu, Q. Liu, H. Xiong, and E. Chen, "A survey on large language models for recommendation," 2024. [Online]. Available: https://arxiv.org/abs/2305.19860

[3] L. Friedman, S. Ahuja, D. Allen, Z. Tan, H. Sidahmed, C. Long, J. Xie, G. Schubiner, A. Patel, H. Lara, B. Chu, Z. Chen, and M. Tiwari, "Leveraging large language models in conversational recommender systems," 2023. [Online]. Available: https://arxiv.org/abs/2305.07961

[4] S. Yun and Y.-k. Lim, "User experience with llm-powered conversational recommendation systems: A case of music recommendation," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, ser. CHI '25. New York, NY, USA: Association for Computing Machinery, 2025. [Online]. Available: https://doi.org/10.1145/3706598.3713347

[5] M. Brucato, A. Abouzied, and A. Meliou, "Package queries: efficient and scalable computation of high-order constraints," *The VLDB Journal*, vol. 27, no. 5, p. 693–718, Oct. 2018. [Online]. Available: https://doi.org/10.1007/s00778-017-0483-4

[6] A. Fariha, S. Roy, and A. Meliou, "Sudocu: Summarizing documents by example," in *Proceedings of the VLDB Endowment*, vol. 13, no. 12, 2020, pp. 2861–2864. [Online]. Available: https://www.vldb.org/pvldb/vol13/p2861-fariha.pdf

[7] S. Narayan *et al.*, "Subsume: A dataset for subjective summary extraction from wikipedia documents," in *Proceedings of the 1st NewSumm Workshop*, 2021. [Online]. Available: https://aclanthology.org/2021.newsum-1.14.pdf