Asymmetric Personalization on LaMP-QA

Ozel Yilmazel, James Allan, Hamed Zamani

UMassAmherst
College of Information & Computer Sciences
Center for Intelligent Information Retrieval

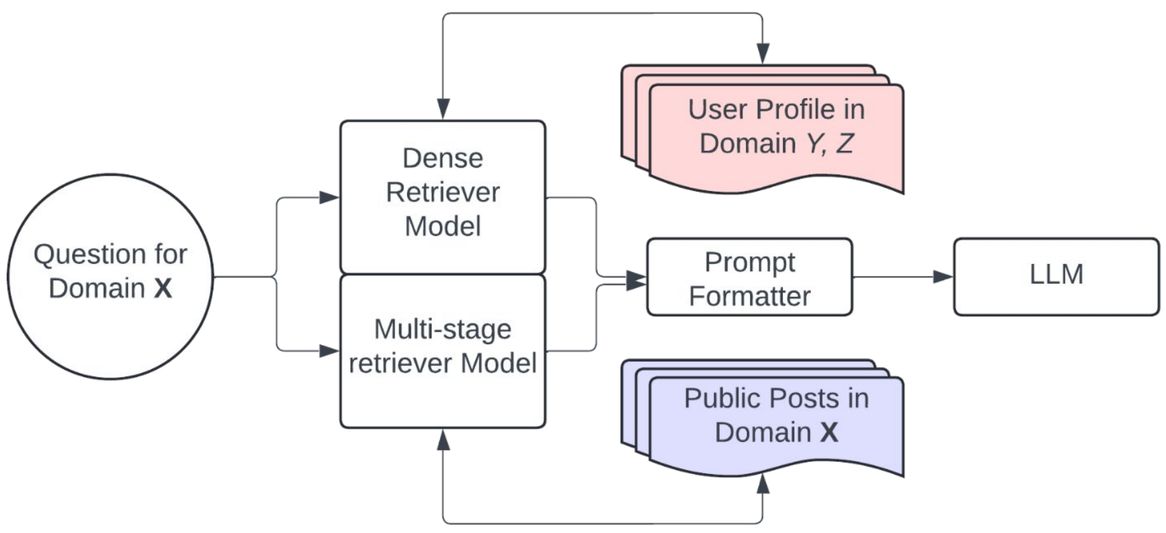
Introduction

Most personalization with LLMs focuses on *symmetric* settings, where generated content matches the domain of user input. We instead explore *asymmetric* personalization, where same-domain data is unavailable. **By combining public domain knowledge with users' out-of-domain inputs, we achieve stronger, more grounded, and factually accurate results while preserving personalization.** Using LaMP-QA avoided costly data collection and revealed that users' writing styles and information needs vary across three broad topic categories, providing a natural testbed for asymmetric personalization.

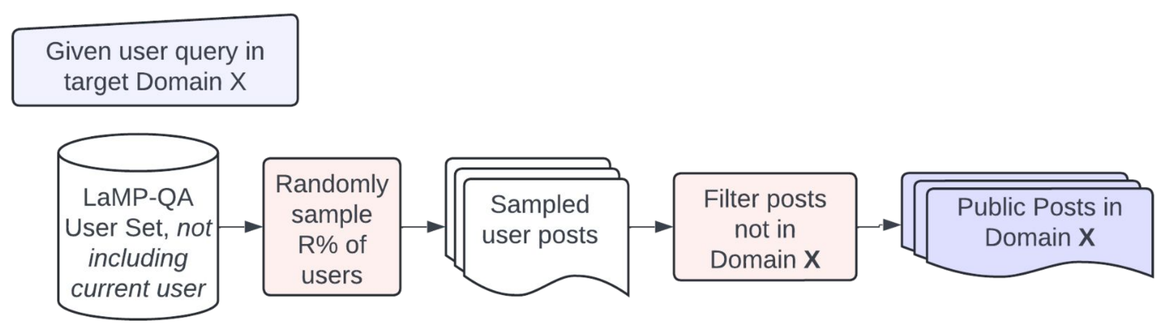
Methods

We compare several methods: **NoPers**, a regular chatbot with no user data; **RAG-Full**, which uses all user data including the target domain; **RAG-Asym**, limited to out-of-domain user data; **2-Aug-Asym**, which combines out-of-domain user data with public in-domain posts; **2-Aug-Asym-Cat**, which adds category-based prompting for relevance; and **2-Aug-Full**, which extends 2-Aug with access to all user data.

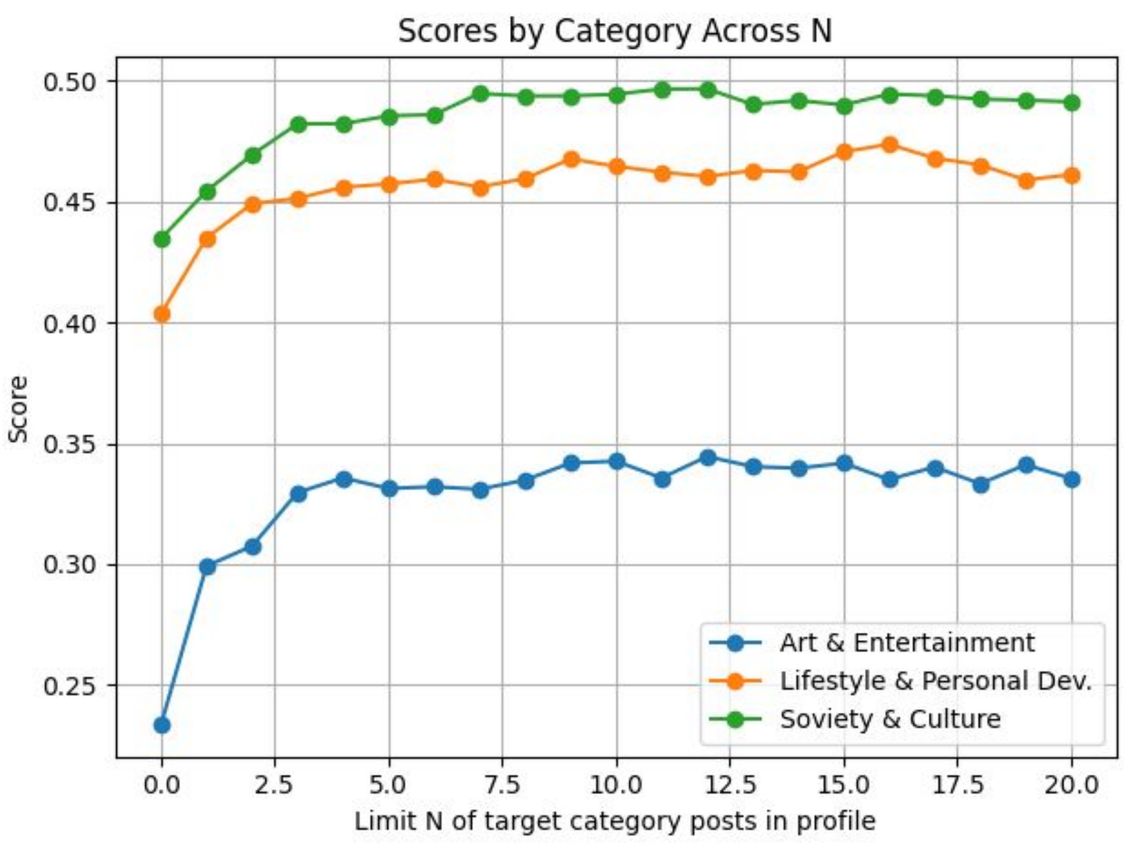
2-Aug-Asym



Public Sampling



Results



Method	*A&E*	*L&P*	*S&C*	*Average*
NoPers	0.2931	0.4658	0.4748	0.4112
RAG-Asym	0.2335	0.4035	0.4181	0.3517
2-Aug-Asym	0.3094	**0.4899**	0.4884	0.4292
2-Aug-Asym-Categorized	**0.3326**	0.4822	**0.4924**	**0.4357**
RAG-Full	0.3321	0.4520	0.4885	0.4242
2-Aug-Full	**0.3195**	**0.4888**	**0.5180**	**0.4421**

Measure	*A&E*	*L&P*	*S&C*
P@5	0.8444	0.6901	0.7708
P@10	0.8284	0.6266	0.7351
P@20	0.7981	0.5419	0.6774

Findings

- **Out-of-domain hurts**: Using only out-of-domain user data often degrades performance—sometimes worse than no personalization. Manual inspection of RAG-Asym outputs revealed excessive noise and spurious cross-domain connections.
- **Retrievers are effective**: Across all domains, ≥50% of top-k retrieved posts are in-domain. RAG-Asym wastes this by ignoring them, while RAG-Full benefits from them.
- **2-Aug restores balance**: By adding even a small amount of public in-domain knowledge, answers become more grounded and relevant. In fact, inserting just one target-domain post noticeably boosts performance.
- **Beyond cold-start**: 2-Aug with full profiles also performs exceptionally well, showing that public knowledge integration strengthens personalization even when user data is abundant.
- **Takeaway:** Combining public knowledge with user features—plus careful prompting—produces stronger, more reliable, and more grounded personalization than either strategy alone.

Future Work

Future work will extend planner-based methods (e.g., PlanPers) to the 2-Aug setting, explore iterative personalization that gradually integrates user data, and develop dynamic model selectors to choose among methods. We also aim to improve retrieval efficiency, test across different LLM families, and build new benchmarks focused on asymmetric personalization. Extending asymmetric research by creating a Tweet/Abstract collection of CS Faculty.