

Specialized Project

KG-DCR: A Knowledge Graph Approach for Diverse Conversational Recommendations

Supervisors:

- *Assoc. Prof. THOẠI NAM*
- *PhD. DIỆP THANH ĐĂNG*

Students:

- **Nguyễn Ngọc Song Thương – 2252803**
- **Nguyễn Hữu Thanh – 2252742**
- **Nguyễn Đức Tâm – 2252734**

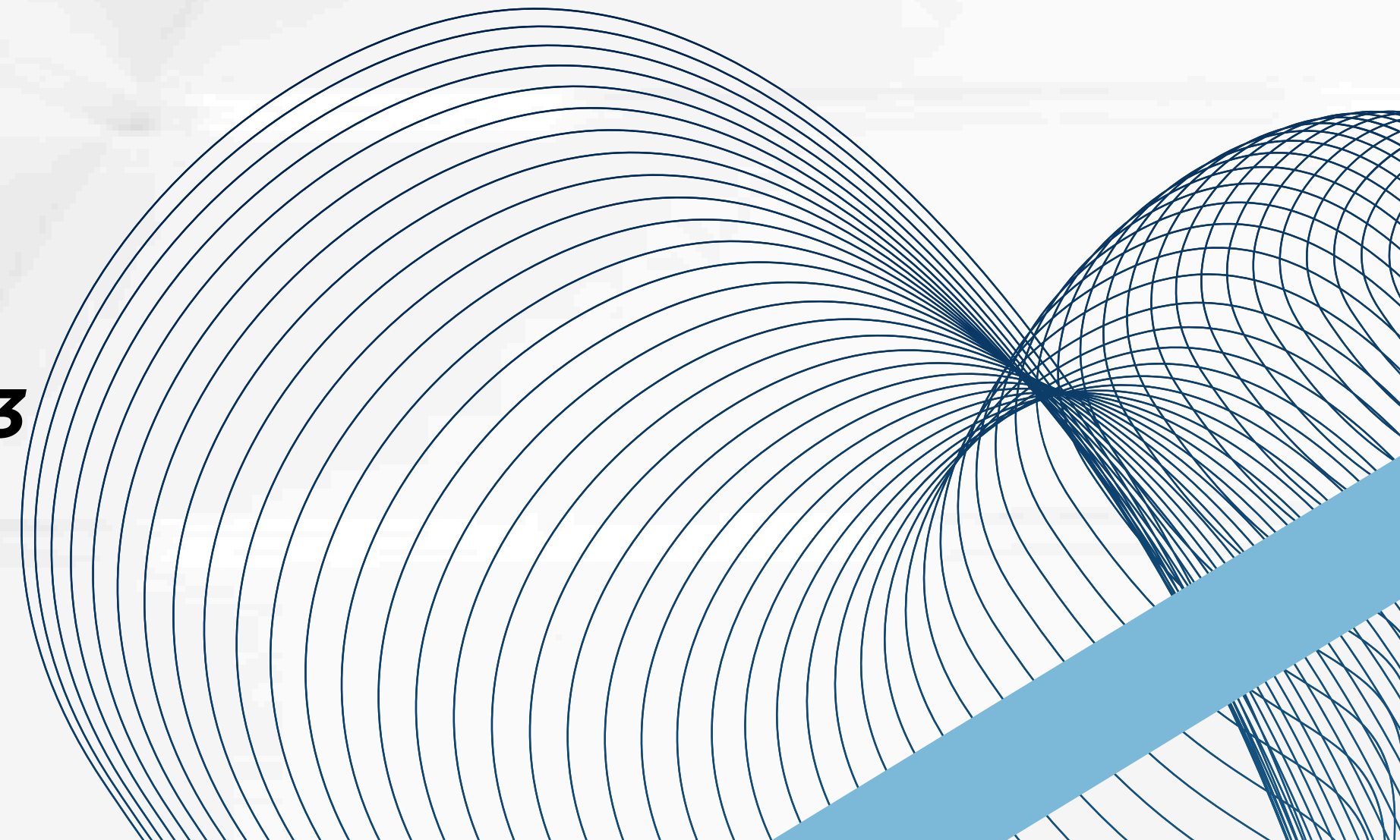


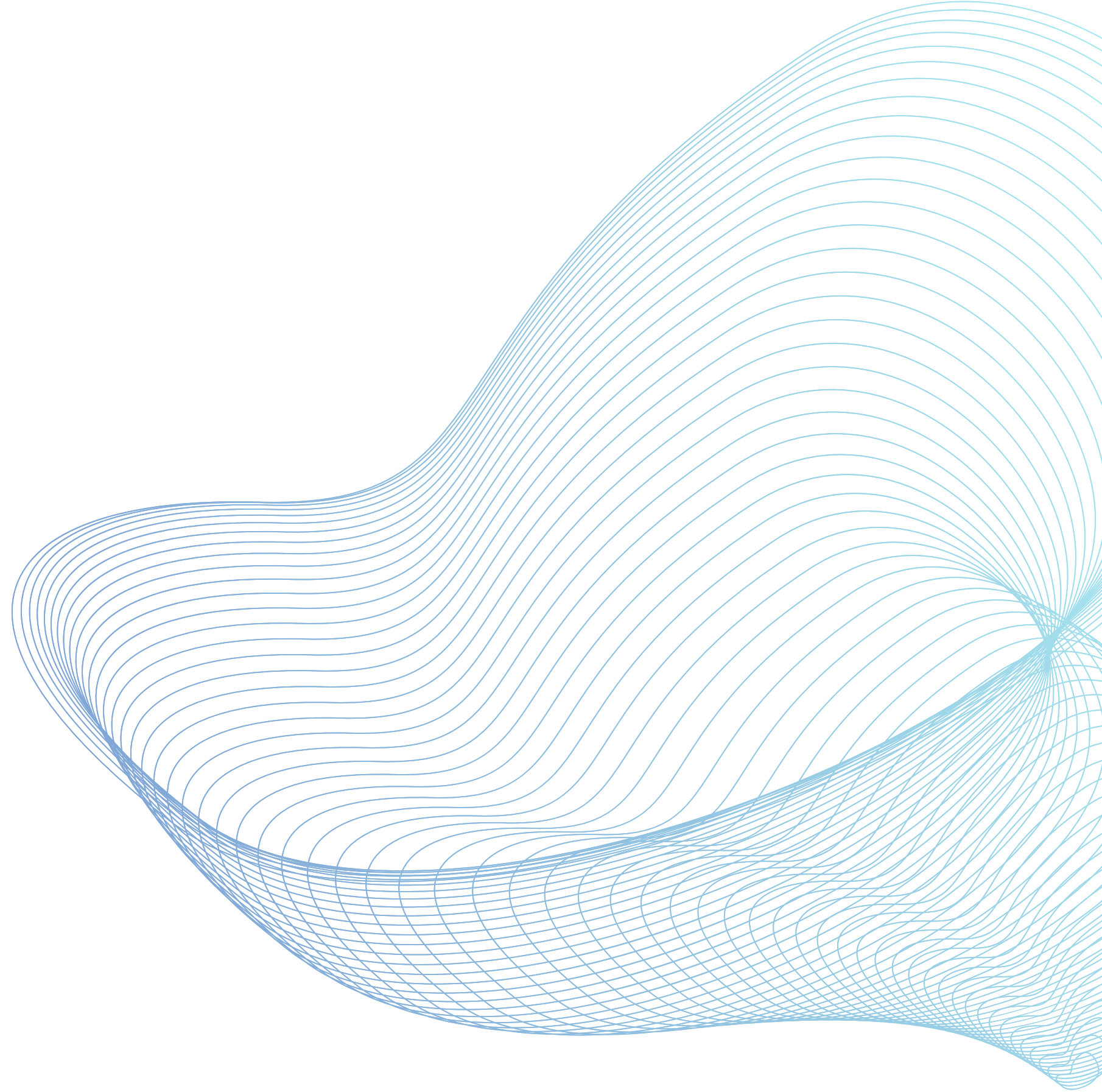
Table of Content

- 1 Introduction & Motivation**
- 2 Background**
- 3 Related Works**
- 4 Metrics/ Datasets**
- 5 Experiments**
- 6 Conclusion & Future Work**



1

Introduction & Motivation



Introduction

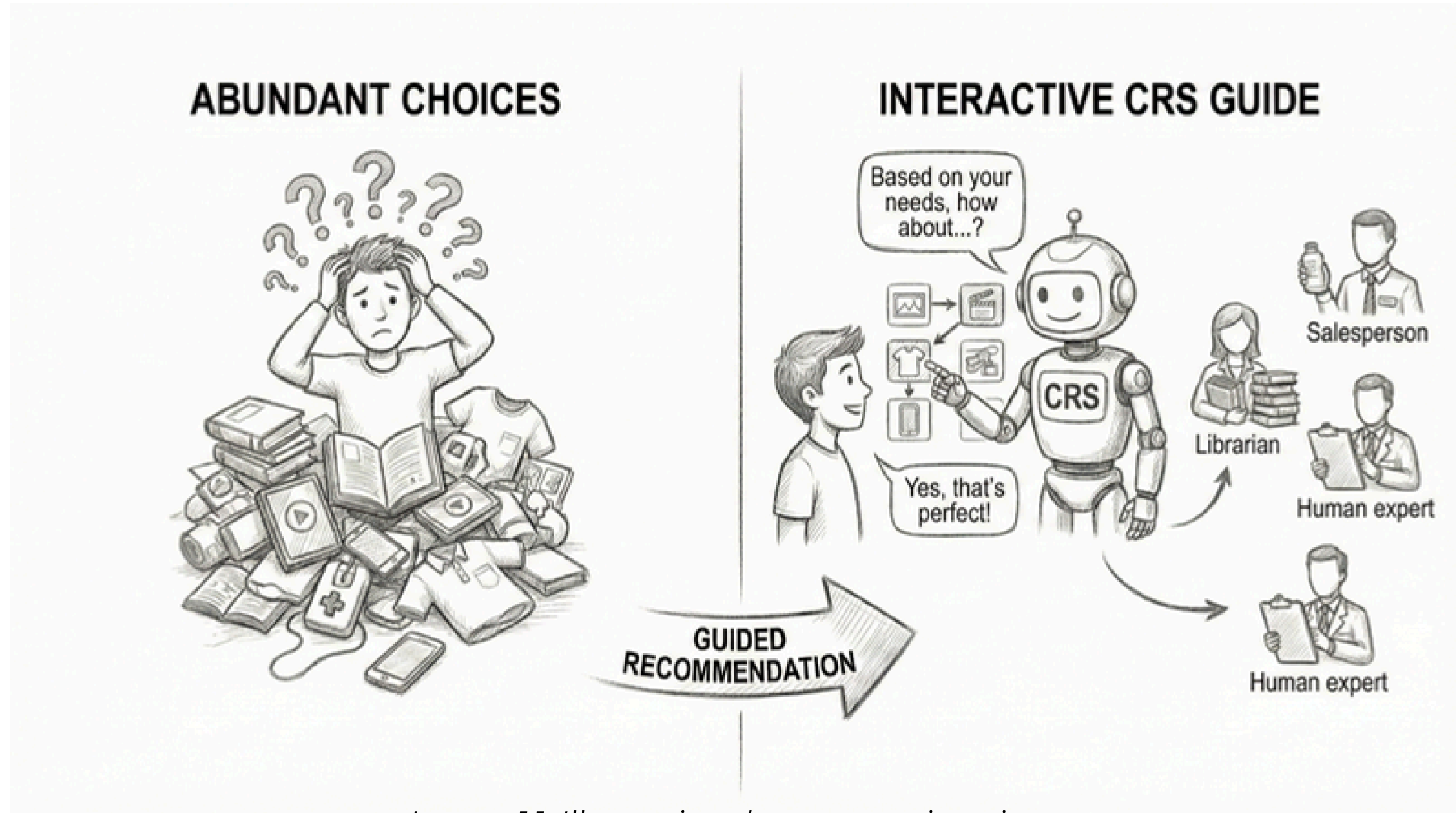


Image 1.1: Illustration the current situation



Image ref: image is generated by Gemini Banana

Motivations

The Accuracy-Diversity Tradeoff Problem

- Encounters only information and opinions that conform to and reinforce their own beliefs
- Cause by algorithms that personalize an individual's online experience.

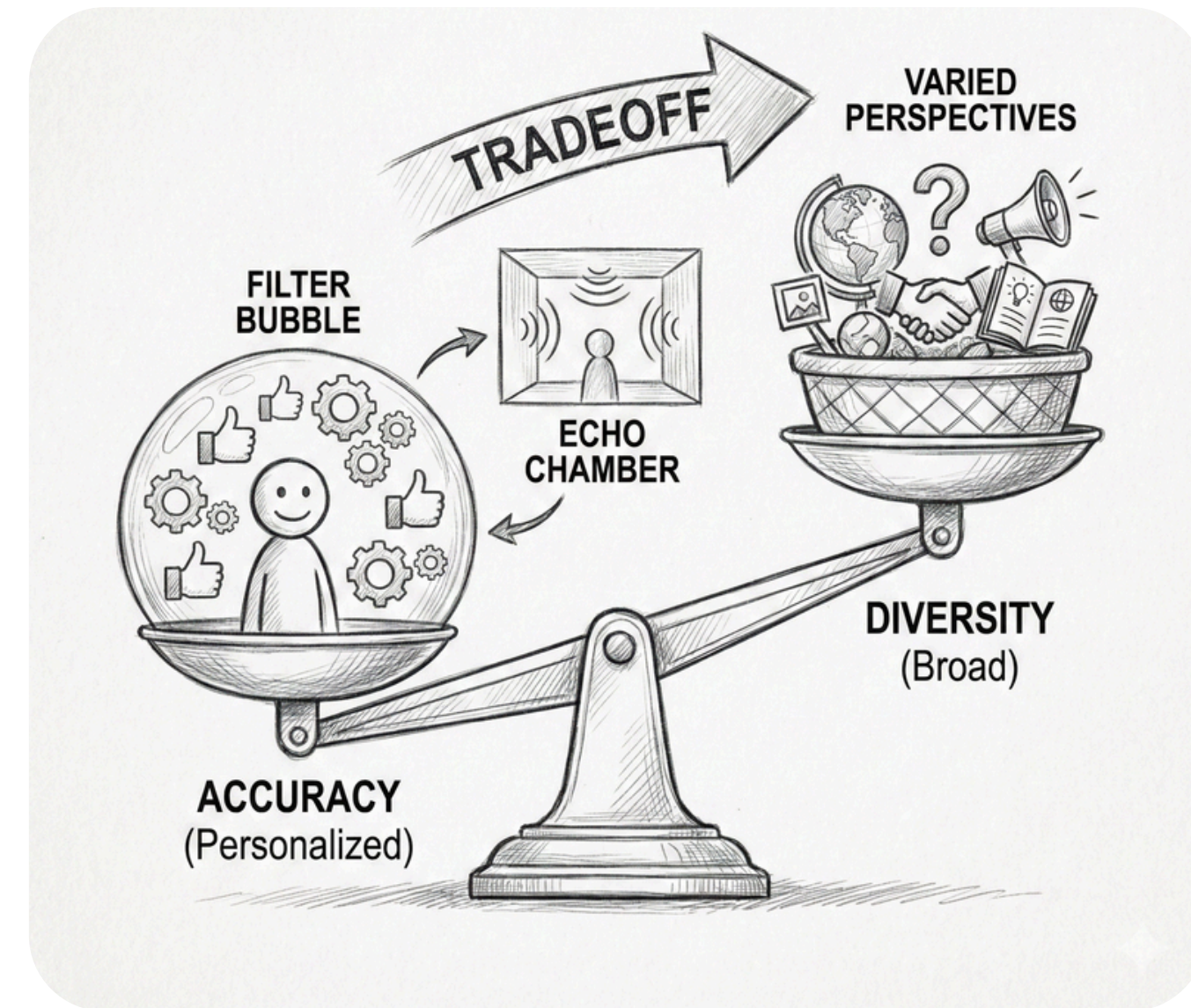


Image 1.2: Illustration the Accuracy-Diversity Tradeoff Problem



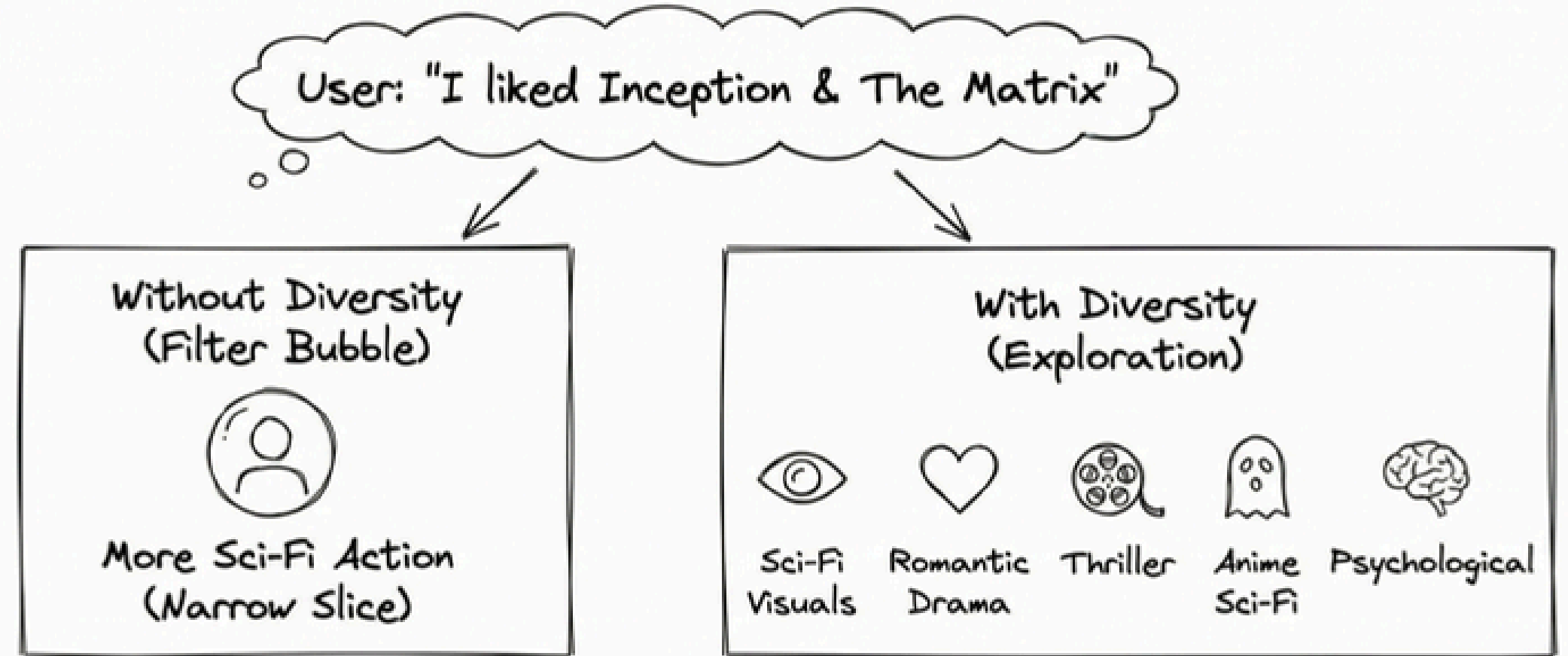
Image ref: image is generated by Gemini Banana

Motivations

The Accuracy-Diversity Tradeoff Problem

Example 1: Movie Recommendation

Movie Recommendation Example: Diversity vs. Filter Bubble



Why It Matters: Diversity -> Exploration -> Productive Conversation

Image 1.3: Illustration the problem about movie recommendation



Image ref: image is generated by Gemini Banana

Motivations

The Accuracy-Diversity Tradeoff Problem

Example 2: News Article Recommendation

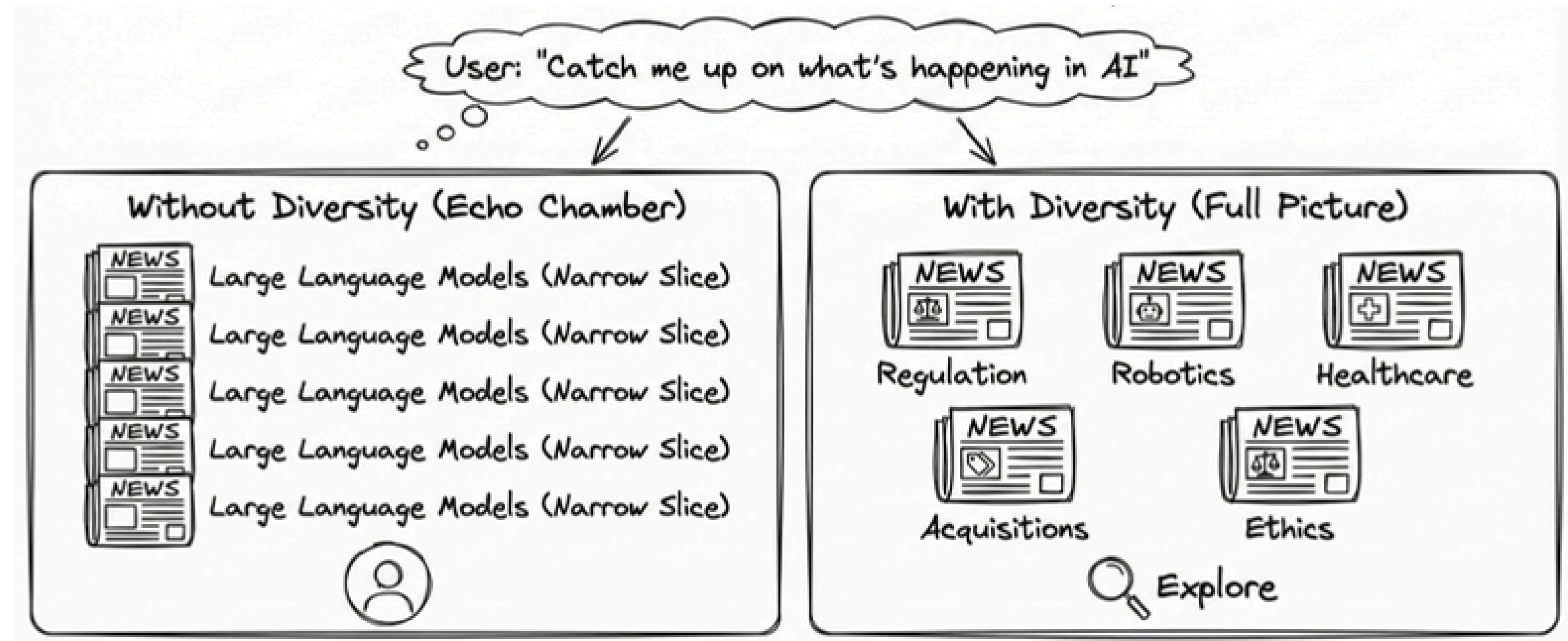


Image 1.4: Illustration the problem about News Article Recommendation



Image ref: image is generated by Gemini Banana

Objectives

For this thesis proposal

- ❗ Investigate **existing techniques** that
 1. Enhance diversity in CRS
 2. Utilize KG for accuracy and diversity in CRS
 3. Utilize LM for response generation in CRS
- ❗ **Comparison and evaluation** of existing techniques
- ❗ Determine the **research direction** for final thesis



Research Questions

RQ1: Diversity via Question Elicitation

How can question elicitation promote diverse preference exploration?

RQ2: KG-Grounded Conversational Recommendation

How to integrate KG-retrieved information with language model for coherent, grounded CRS?

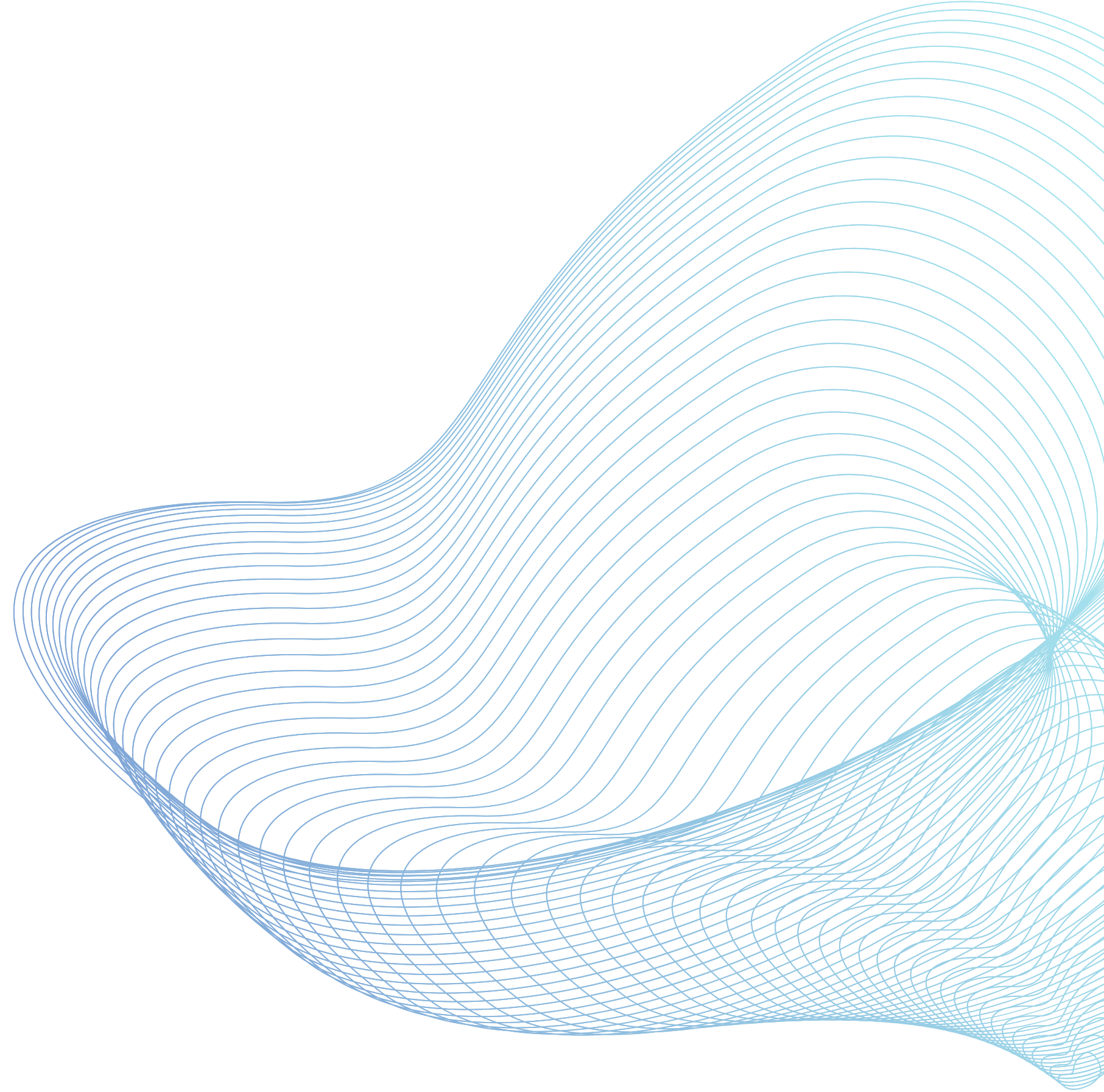
RQ3: Retrieval vs Elicitation Diversity Effects

How do retrieval-level and elicitation-level diversity interact for overall diversity?



2

Background



Conversational Recommender System

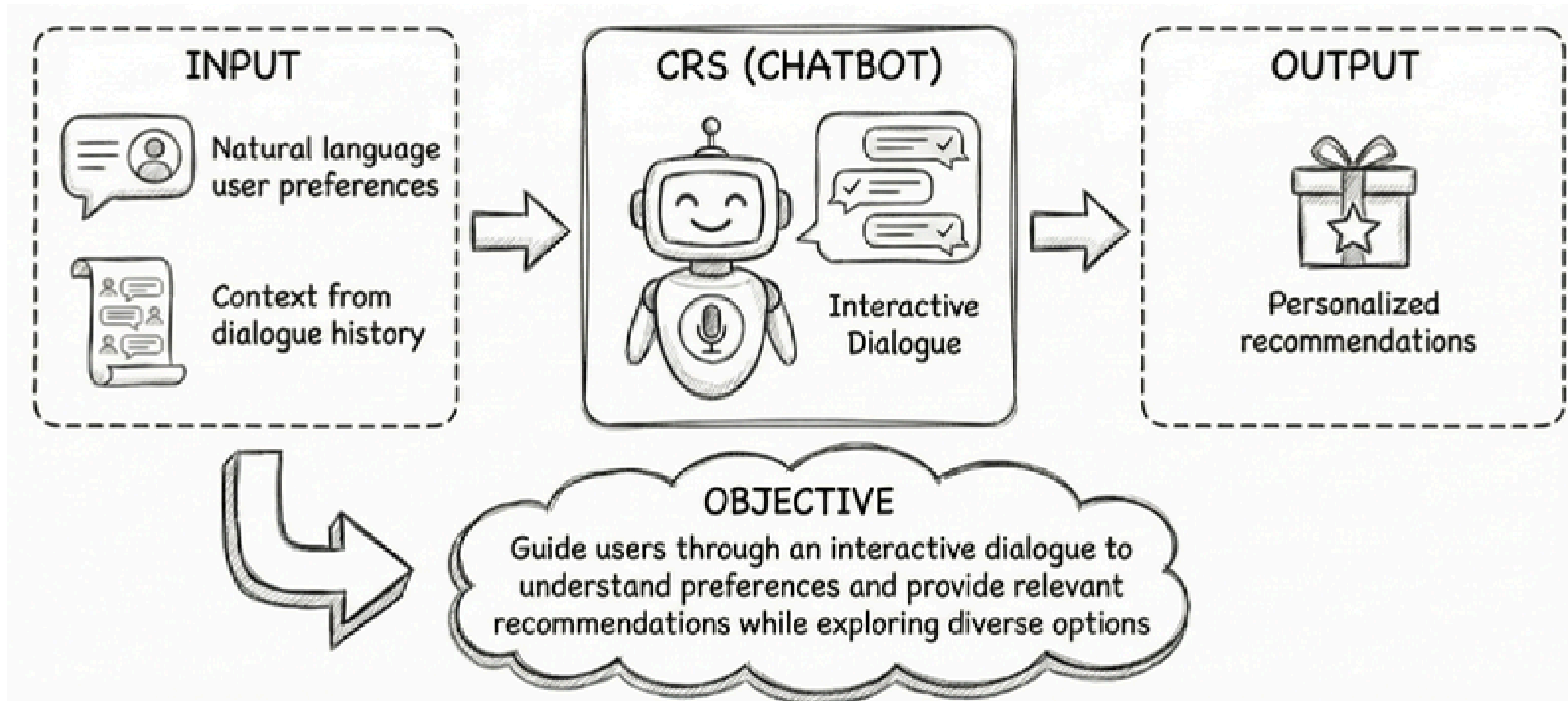


Image 2.1: Illustration the Conversational Recommender System



Image ref: image is generated by Gemini Banana

Knowledge Graph

Definition: A structured representation of entities and their relationships.

Core Roles in CRS:

- Capture rich semantic information about items
- Enabling multi-hop reasoning and relationship-based retrieval

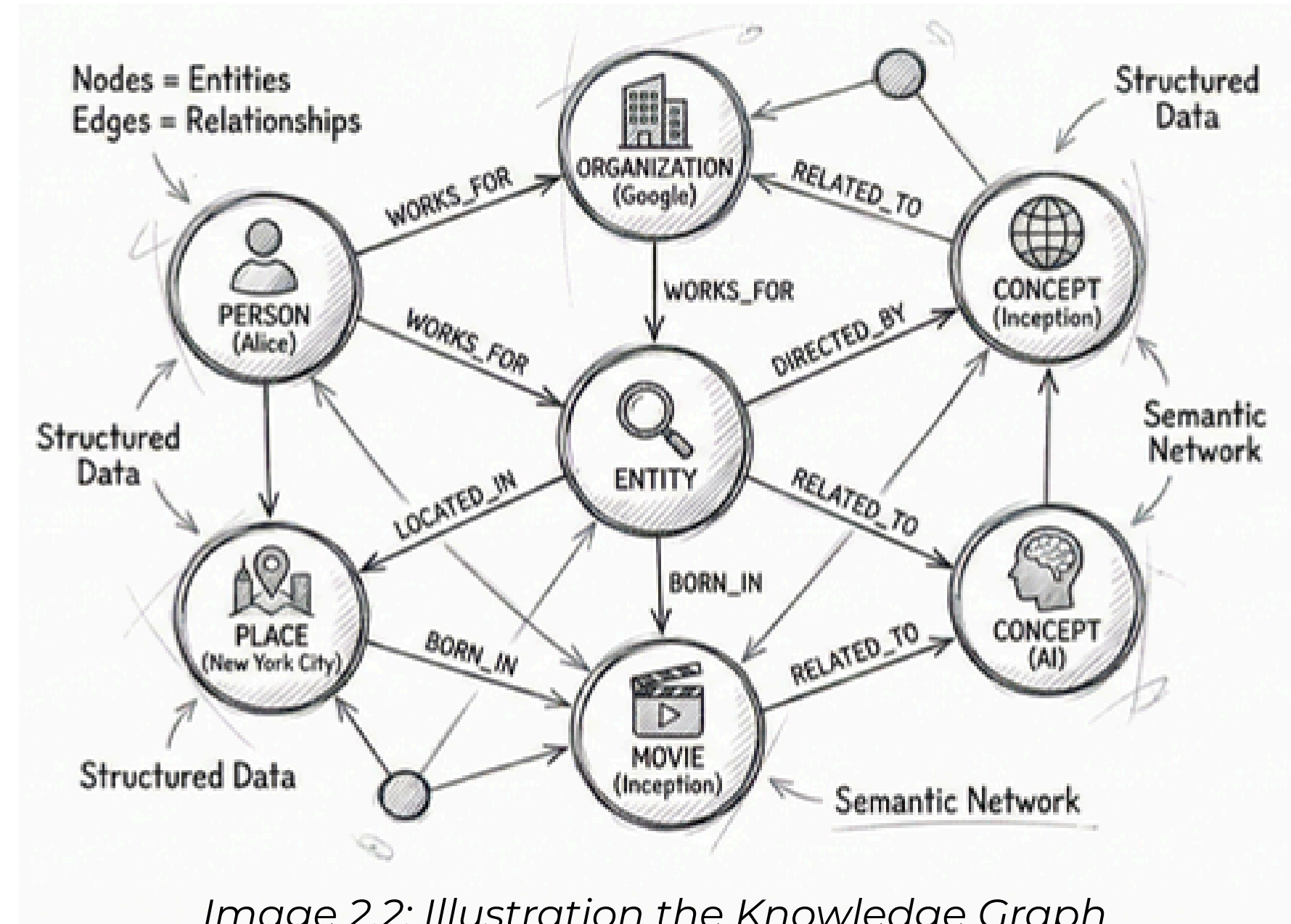


Image ref: image is generated by Gemini Banana

Language Models

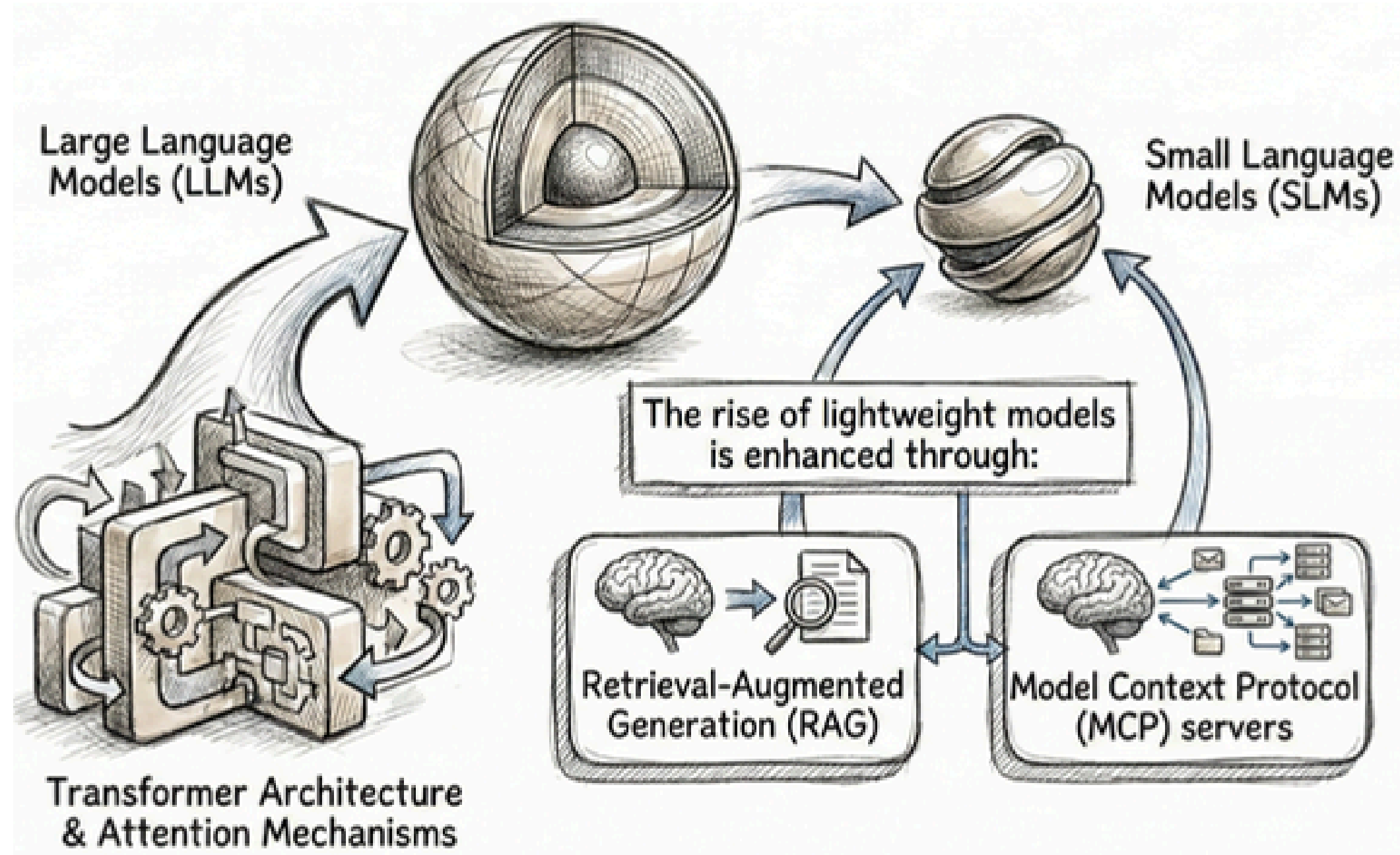


Image 2.3: Illustration the Language Models



Image ref: image is generated by Gemini Banana

3

Related Works



Summary of existing techniques

CRS Paradigms			
Aspects	Objectives	Techniques	Representative Methods
Generation-Based	User interaction quality, engaging chitchating while recommending	Natural dialogue on human-written corpora; subgoal prediction;	KBRD [1], KERS [2], KERL [3], KE CRS [4], CR-Walker [5], ConKD [9]
Attribute-Based	Get to final recommendation with as least turn as possible	Reinforcement learning on simulated sessions;	UNICORN [10], DECREC [11]

Summary of existing techniques

Knowledge Graph Utilization		
Aspects	Techniques	Representative Methods
Representation Learning	KG embeddings (TransE, R-GCN); PLM fusion; graph entity captioning	KERL [3], KBRD [1], COMPASS [7], UNICORN [10], DECREC [11]
Reasoning & Retrieval	Tree-structured reasoning; knowledge gates; GraphRAG with PageRank	CR-Walker [5], KERS [2], G-CRS [8], CAFE [12]

Summary of existing techniques

Language Model Utilization

Aspects	Techniques	Representative Methods
Dialogue Generation	Encoder-decoder; knowledge-enhanced decoders; Teacher-student distillation	BARCOR [13], ConKD [9], KBRD [1]
Token Specialization	Item tokens; placeholder tokens; intent tokens	KECRS [4], KERL [3], ConKD [9]
LLM Reasoning	Knowledge-aware instruction tuning; in-context learning; multi-agent simulation	COMPASS [7], G-CRS [8], KGLA [6]

Summary of existing techniques

Diversity Enhancement		
Aspects	Techniques	Representative Methods
Data-Level	Balanced data simulation; experience replay stratification	LOT-CRS [14], DECREC [11]
Reasoning-Based	Multi-path exploration; discriminative KG rationales	CR-Walker [5], KGLA [6]
Training Objectives	Soft target learning; BOE loss for neighbor mentions	ConKD [9], KE CRS [4]

Diversity-enhanced conversational recommendation via multi-agent reinforcement learning (DECREC)

Core mechanism: Reinforcement Learning

- A triple-agent framework:
 - **Query Action Agent:** decide optimal query to ask user for preference elicitation
 - **Recommend Action Agent:** decide optimal items for recommendation
 - **Decision Agent:** decide whether to ask or recommend

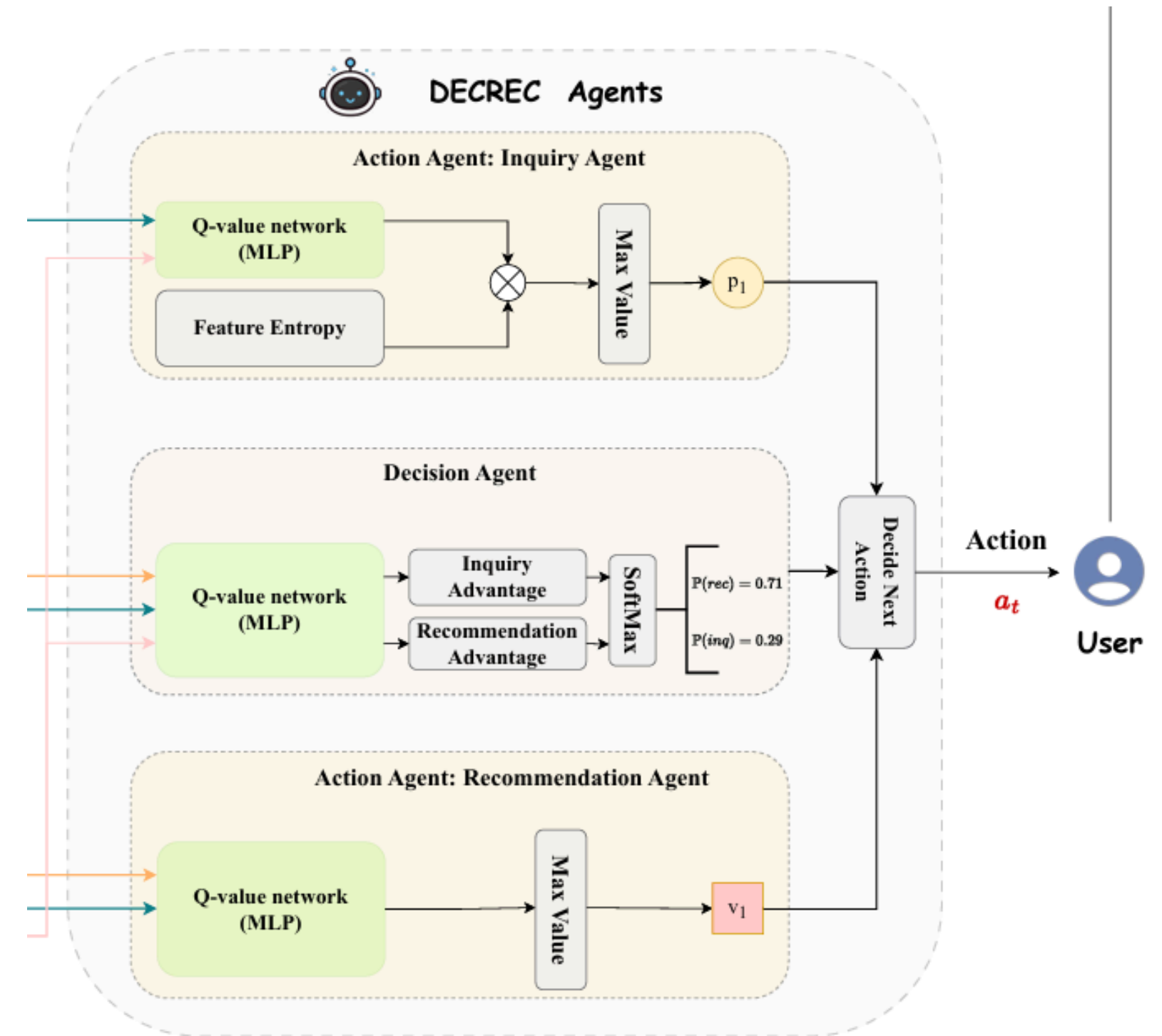


Figure 3.1: Agents Component of DECREC [11]

Diversity-enhanced conversational recommendation via multi-agent reinforcement learning (DECREC)

How DECREC enhances diversity

- **Dynamically Calibrated Experience Replay (DCER):**
 - Address "long-tail" issue - frequent interactions with popular items and ignore"
 - Balanced sampling of head action pool and long tail action pool
- **Feature Entropy:**
 - Calculate with the probability a feature appear in all candidate items
 - Prompt Agent to select high-entropy features for more extensive exploration.

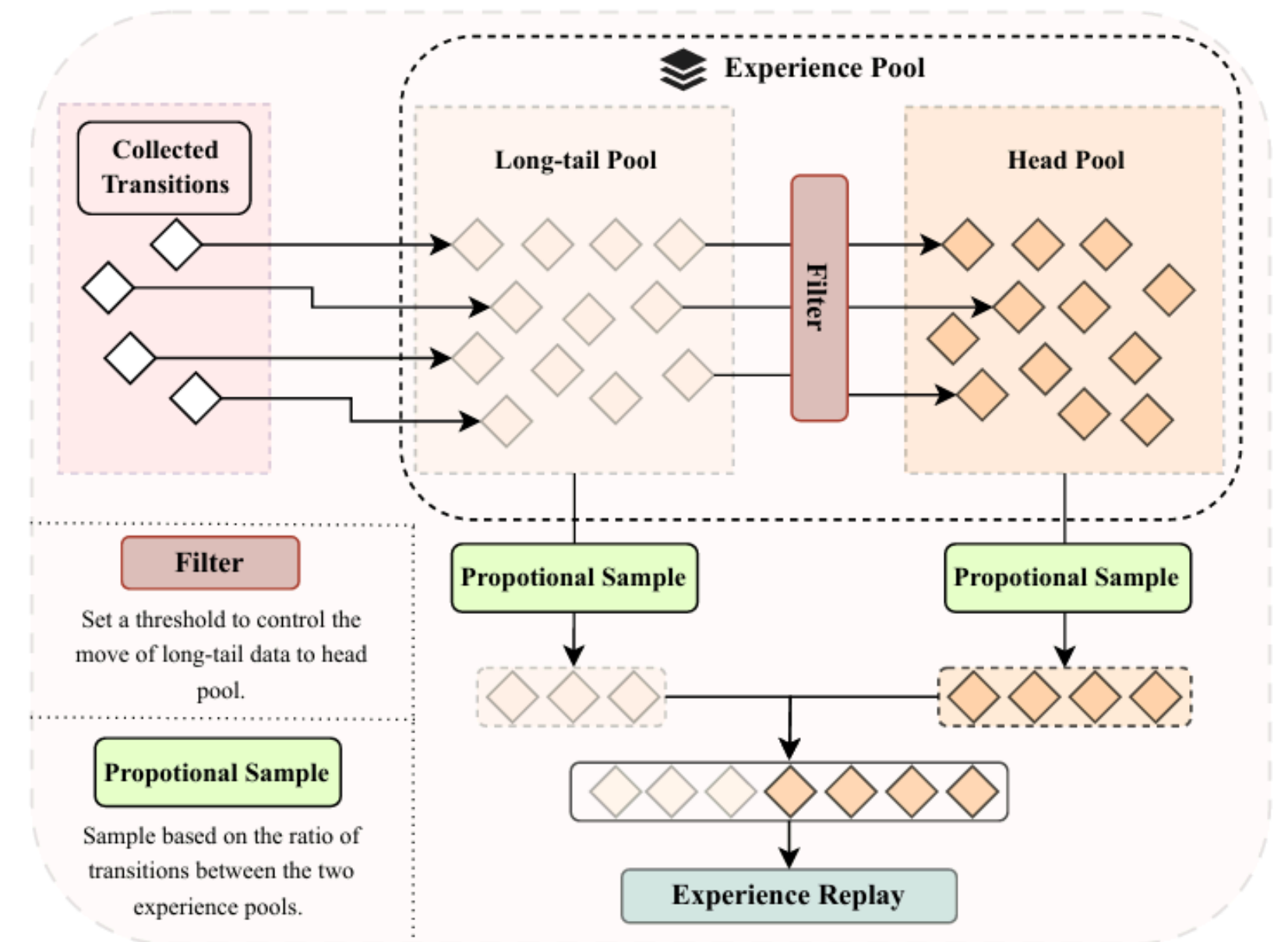


Figure 3.2: The procedure of DCER [11]

Diversity-enhanced conversational recommendation via multi-agent reinforcement learning (DECREC)

Utilization of Knowledge Graph

- **Local Short-term Graph**
 - Immediate interactions of user, items, and features in current dialogue
 - Learn short-term user, items, and features embedding to capture immediate interest (GCN)
- **Global Long-term Graph**
 - All interactions of all users, items, and features
 - Learn historical interests and global embeddings of users, items, and features

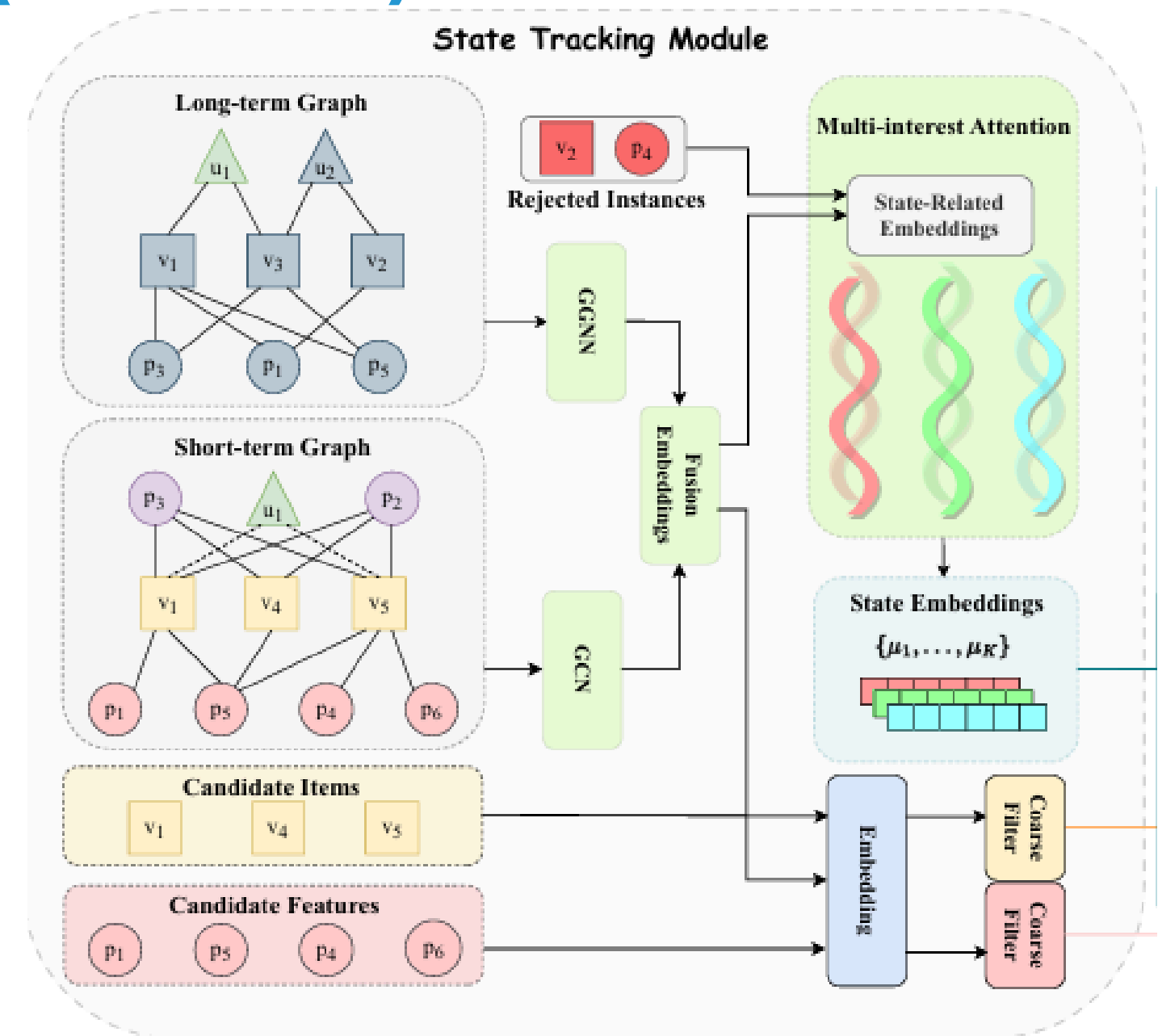


Figure 3.3: State Tracking Module of DECREC [11]

Tree-Structured Graph Reasoning and Dialog Acts for Conversational Recommendation (CR-Walker)

Mechanism: Multi-path reasoning

- Simple 3-way classification to decide **system intent (recommend, query, chat)**
- Rule-based narrowing of search space for each intent type

Intent	Hop 1	Hop 2
Recommend	mentioned attributes + mentioned items' attributes	candidate items
Query	generic classes	attributes
Chat	mentioned entities	all entities

Table 3.1: Reasoning rules for narrowing down CRWalker's search space [5]

Tree-Structured Graph Reasoning and Dialog Acts for Conversational Recommendation (CR-Walker)

Mechanism: Multi-path reasoning

• Entity selection - Walker Cell

- Input: utterance embedding, portrait embedding, selected intent embedding
- Gated mechanism => **context embedding**
- **Selection scores** are the similarity scores between context embedding and candidate entity embedding (sigmoid of dot product)
- Selection criteria: **selection scores** \geq **threshold**, restricted by maximum number of entities on each hop

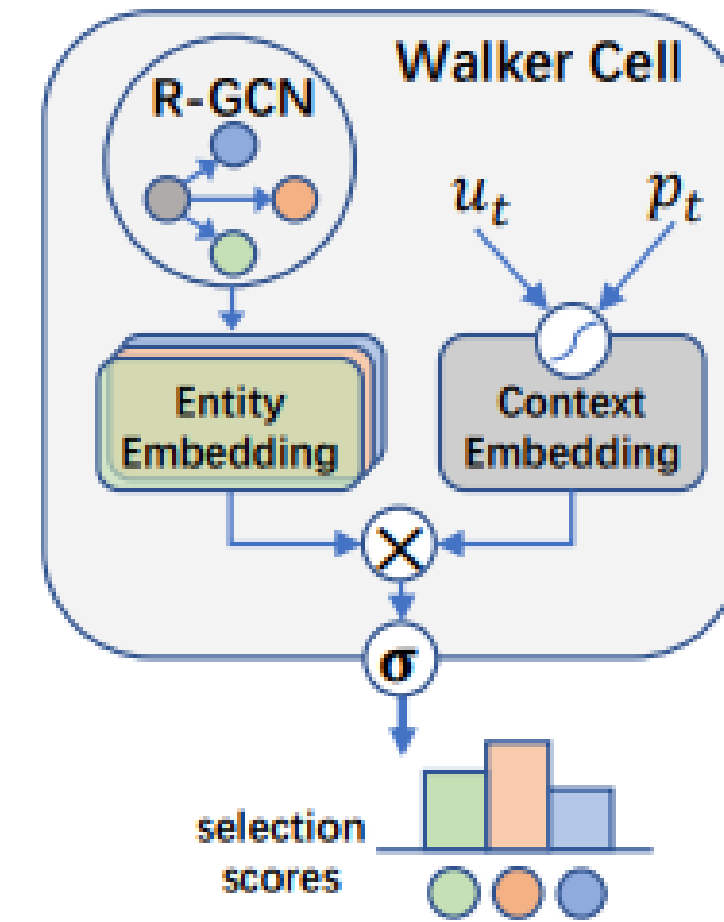


Figure 3.4: Walker Cell in CRWalk [5]

Tree-Structured Graph Reasoning and Dialog Acts for Conversational Recommendation (CR-Walker)

Mechanism: **Multi-path reasoning**

- **Dialog Act:**

- Preorder traversal of the final selected tree
- Input for GPT-2 to generate response to user

Strengths

- ✓ Explore multipath => increased diversity

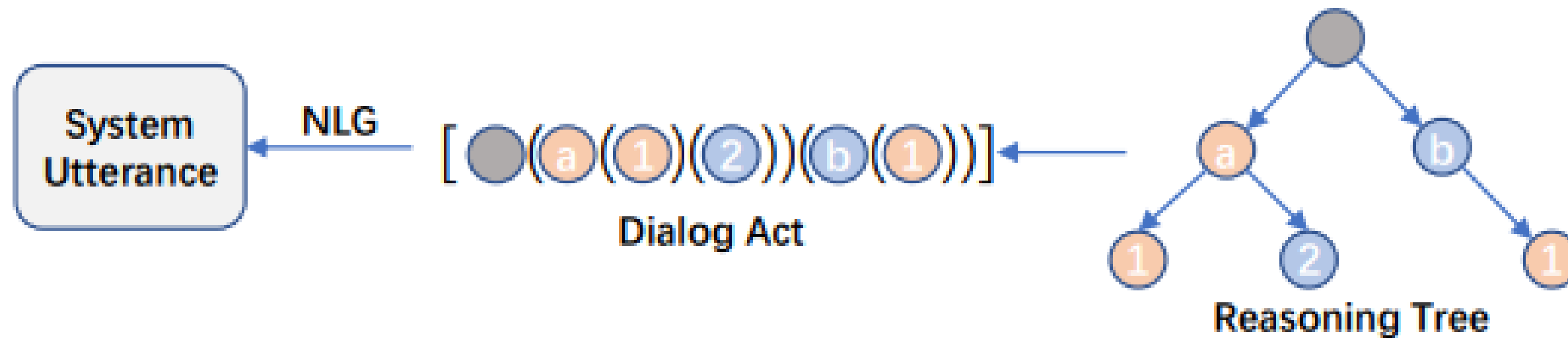


Figure 3.5: The formation of dialog act in CRWalker [5]

Retrieval-Augmented Generation with Graphs (GraphRAG)

Core mechanism: Retrieval-Augmented Generation

- A **training-free** framework build upon Graph Retrieval for augmented information
- The knowledge graph captures the relationship of **conversation-entity interaction**

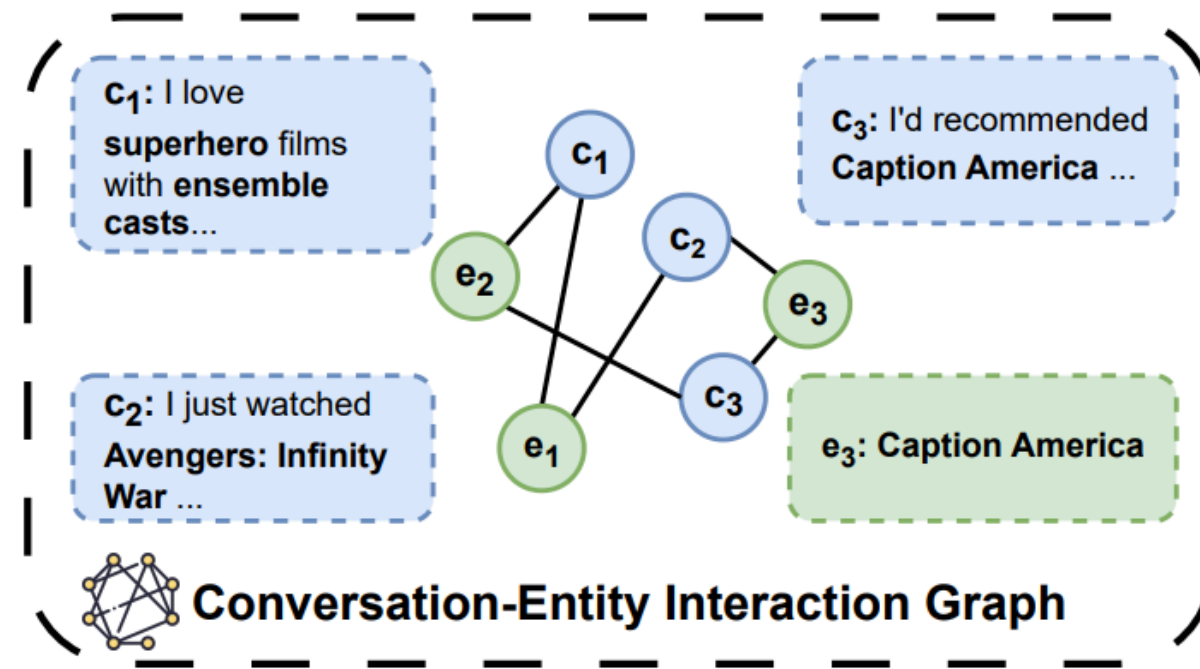


Figure 3.6: Illustration of conversation-entity interaction graph [8]

Retrieval-Augmented Generation with Graphs (GraphRAG)

Core mechanism: Retrieval-Augmented Generation

- Retrieval process:
 - **Graph Reasoner** expand entities set from the mentioned entities
 - Using **Personalized Page Rank** to retrieve most relevant conversation examples and candidate entities

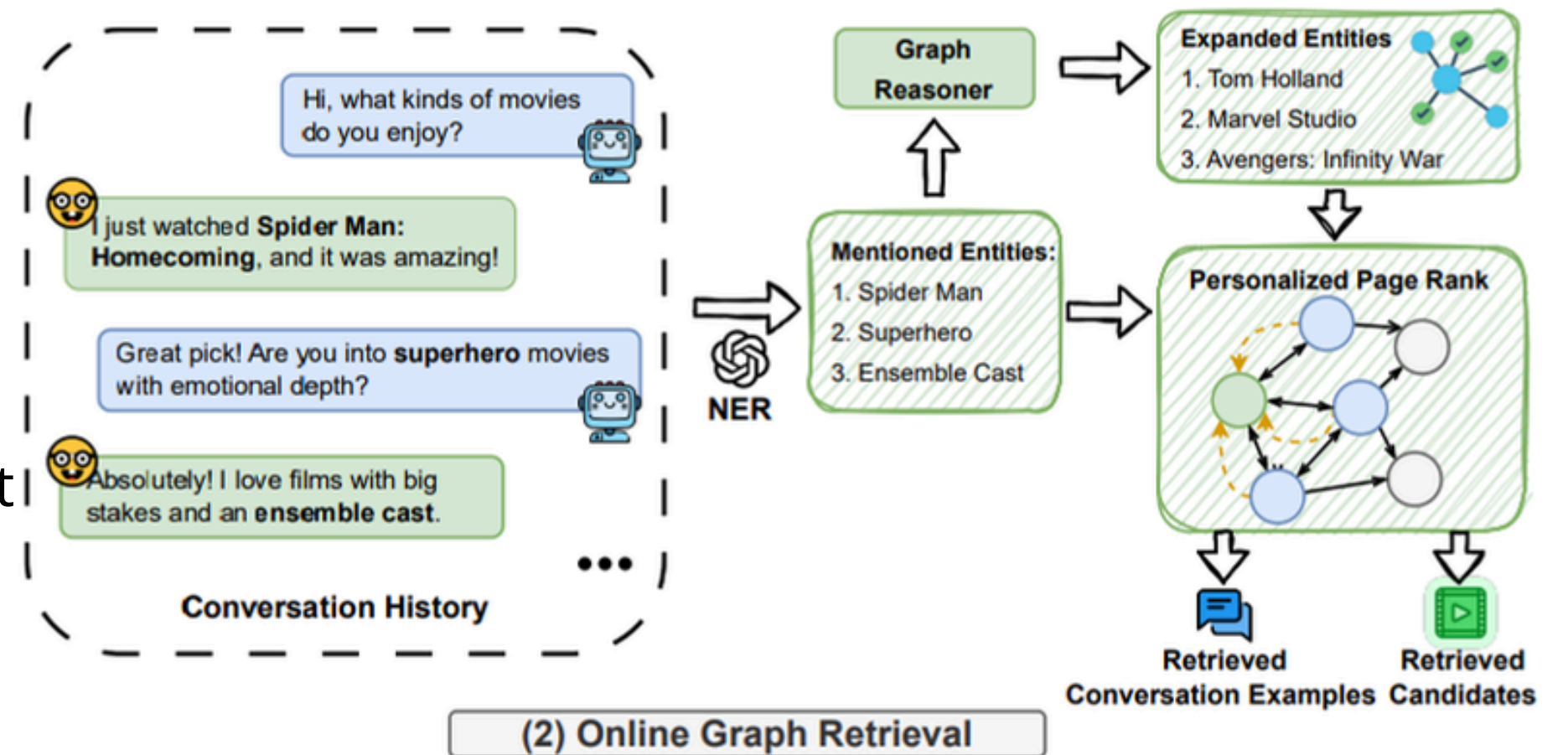


Figure 3.7: Online Graph Retrieval Process [8]

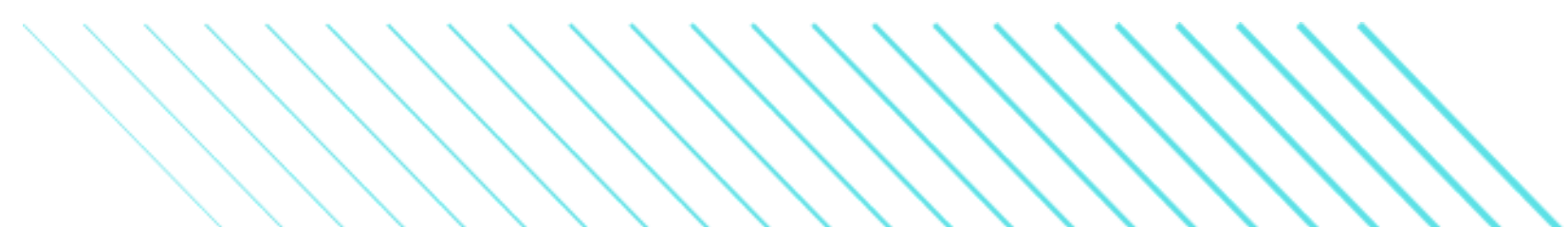
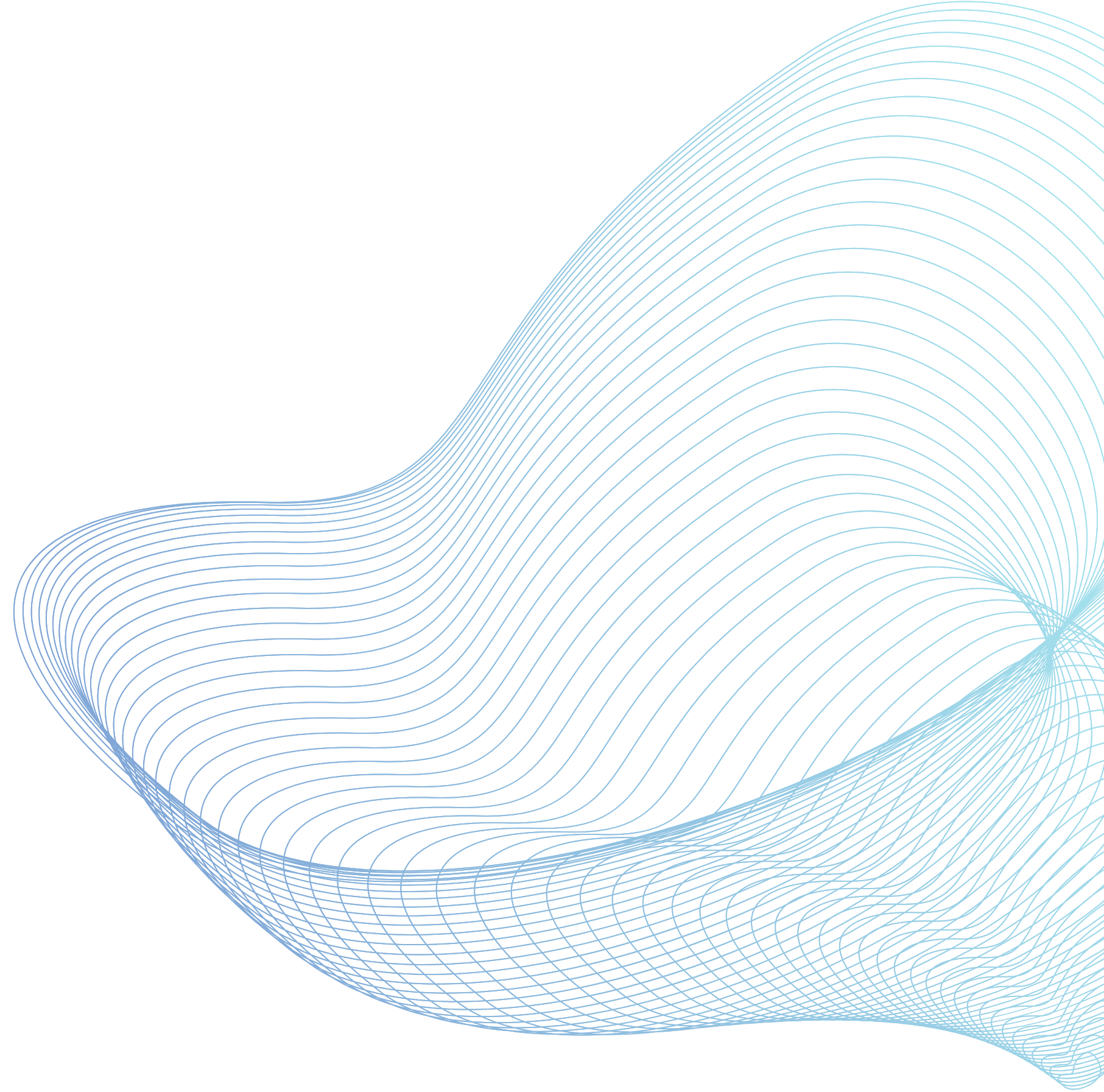
Strengths

- ✓ No fine-tuning needed
- ✓ Community examples => Avoid cold start



4

Metrics/ Datasets



CRS Metrics

General Recommendation Metrics

Hit Rate at K (HR@K):

Whether the recommender retrieves at least one relevant item in the Top-K list.

$$\text{HR@K} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(|h(i, K) \cap R_i| \neq \emptyset)$$

Mean Reciprocal Rank (MRR@K):

Focuses on the rank of the first relevant item.

$$\text{MRR@K} = \frac{1}{N} \sum_{i=1}^N \begin{cases} \frac{1}{\min_{v \in R_i} \text{rank}(v)} & \text{if } \min_{v \in R_i} \text{rank}(v) \leq K \\ 0 & \text{otherwise} \end{cases}$$

Recall at K (Recall@K):

The proportion of all relevant items that appear in the Top-K recommendations.

$$\text{Recall@K} = \frac{1}{N} \sum_{i=1}^N \frac{|L(i, K) \cap R_i|}{|R_i|}$$

Normalized Discounted Cumulative Gain (NDCG@K):

Measures ranking quality by rewarding relevant items placed at higher positions.

$$\text{DCG@K} = \sum_{j=1}^K \frac{2^{\text{rel}_i(j)} - 1}{\log_2(j + 1)} \quad \text{NDCG@K} = \frac{1}{N} \sum_{i=1}^N \frac{\text{DCG}_i@K}{\text{IDCG}_i@K}$$



CRS Metrics

Conversational Recommendation Metrics

Success Rate (SR@T):

Measures whether the system successfully recommends within T turns.

$$SR@T = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(\exists t \leq T, \exists v \in R_i : \text{rank}_{i,t}(v) \leq K)$$

Average Turns (AT):

The success turn for session i as the first turn.

$$AT = \frac{1}{N} \sum_{i=1}^N t_i^*$$

$$t_i^* = \min\{t \in \{1, \dots, T\} : \exists v \in R_i, \text{rank}_{i,t}(v) \leq K\}$$

Hybrid Turn-Rank Metrics (hDCG / hNDCG):

Simultaneously rewarding earlier turns and higher recommendation ranks in a single, unified metric.

$$hDCG@(T, K) = \sum_{t=1}^T \sum_{k=1}^K r(t, k) \times \left[\frac{1}{\log_2(t+2)} + \left(\frac{1}{\log_2(t+1)} - \frac{1}{\log_2(t+2)} \right) \frac{1}{\log_2(k+1)} \right]$$



CRS Metrics

Diversity Metrics

Coverage at the Correct Recommendation Round (COV)

Focuses on the diversity of items surfaced at the time the system “commits” to a final recommendation.

$$\text{COV} = \frac{\left| \bigcup_{s \in \mathcal{S}_{\text{succ}}} L_{t^*(s)}(s) \right|}{|\mathcal{V}|},$$



CRS Dataset

Attribute-based Datasets

Dataset	Domain	Attribute Structure	Used by
LastFM	Music	Artist tags; binary questions	UNICORN [10], DECREC [11]
Yelp	Business	Hierarchical, enumerated questions	UNICORN [10], DECREC [11]
Movie	Film	Genre, director, actor attributes	DECREC [11]
Amazon-Book	E-commerce	Product category and properties	DECREC [11]



CRS Dataset

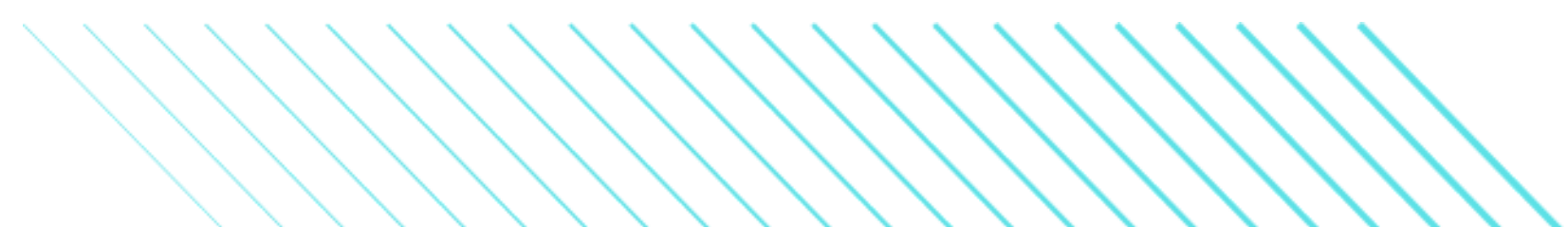
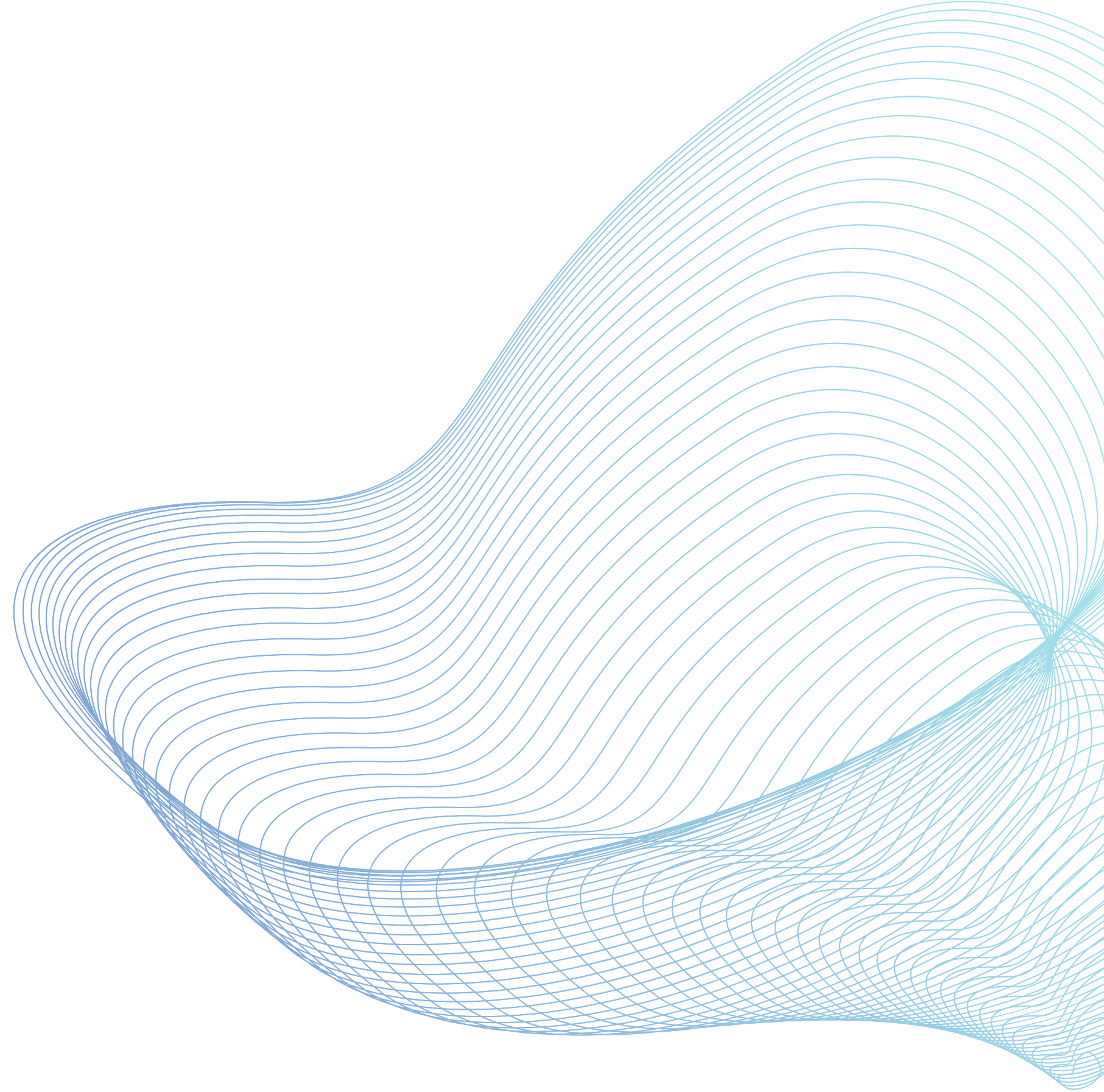
Generation-based Datasets

Dataset	#Dialogues	#Turns	Key characteristics	Used by
ReDial	11,348	~182000	De facto benchmark; natural chit-chat	CR-Walker [5]
GoRecDial	9,125	170,904	Candidate-grounded evaluation	CR-Walker [5]



5

Experiments



Environment Setup

Hardware

- **CPU:** AMD Ryzen 9 9900X (12 cores / 24 threads), base frequency 4.40 GHz
- **RAM:** 64.0 GB
- **GPU:** NVIDIA GeForce RTX 5060 Ti

Deep Learning Stack (GPU)

- **Python :** Python 3.9 / 3.10 / 3.12 (project-dependent), managed via isolated
- **PyTorch:** torch==2.9.1 (stable)
- **CUDA runtime:** CUDA 12.8



Conversation Efficiency Metrics

- **Metrics:** SR@15, AT, hDCG@15
- **Paper:** UNICORN [10], DECREC [11]
- **Datasets:** Amazon-Book, Yelp, LastFM, MovieLens

Paper / Model	Amazon-Book			Yelp		
	SR@15↑	AT↓	hDCG@15↑	SR@15↑	AT↓	hDCG@15↑
UNICORN (paper)	0.547	10.71	0.236	0.412	11.73	0.163
DECREC (paper)	0.594	9.92	0.242	0.457	11.59	0.172
UNICORN (our env)	<u>0.613</u>	9.92	<u>0.264</u>	0.423	10.95	0.182
DECREC (our env)	0.585	<u>9.23</u>	0.251	<u>0.434</u>	<u>10.82</u>	<u>0.192</u>

Table 5.1: Conversation efficiency metrics on MCR benchmarks (AmazonBook, Yelp)



Conversation Efficiency Metrics

- **Metrics:** SR@15, AT, hDCG@15
- **Paper:** UNICORN [10], DECREC [11]
- **Datasets:** Amazon-Book, Yelp, LastFM, MovieLens

Paper / Model	LastFM			MovieLens		
	SR@15↑	AT↓	hDCG@15↑	SR@15↑	AT↓	hDCG@15↑
UNICORN (paper)	0.783	8.11	0.314	0.773	5.51	0.330
DECREC (paper)	0.848	7.31	0.345	0.851	4.69	0.539
UNICORN (our env)	0.801	<u>6.88</u>	0.351	0.799	<u>3.84</u>	<u>0.568</u>
DECREC (our env)	<u>0.845</u>	7.21	<u>0.371</u>	<u>0.814</u>	5.68	0.361

Table 5.2: Conversation efficiency metrics on MCR benchmarks (LastFM, MovieLens)



Conversation Efficiency Metrics

- **Metrics:** Recall@1, Recall@10, Recall@50, T@1, T@3, C@1, C@3
- **Paper:** CR-Walker [5]
- **Datasets:** ReDial, GoRecDial

Model	ReDial			GoRecDial			
	Recall@1↑	Recall@10↑	Recall@50↑	T@1↑	T@3↑	C@1↑	C@3↑
CR-Walker (paper)	0.040	0.187	0.376	0.811	0.975	0.813	0.981
CR-Walker (our env)	0.037	0.154	0.333	0.821	0.976	0.809	0.977

Table 5.3: Conversation efficiency metrics on MCR benchmarks (ReDial, GoRecDial)

Note

- **T@k:** the hit rate among top-k recommendations at each turn [5]
- **C@k:** the hit rate among top-k recommendations only at the end of each dialog [5]



Diversity & Coverage metrics

- **Metrics:** Cov
- **Paper:** DECREC [11], CR-Walker [5]
- **Datasets:** Amazon-Book, Yelp, LastFM, MovieLens, ReDial

Paper / Model	Amazon-Book	Yelp	LastFM	MovieLens	ReDial [*]
	Cov↑	Cov↑	Cov↑	Cov↑	Cov↑
UNICORN (paper)	0.337	0.090	0.700	0.124	-
DECREC (paper)	0.365	0.123	0.743	0.169	-
DECREC (our env)	0.451	0.152	0.969	0.268	-
CR-Walker (paper)	-	-	-	-	0.337
CR-Walker (our env)	-	-	-	-	0.313

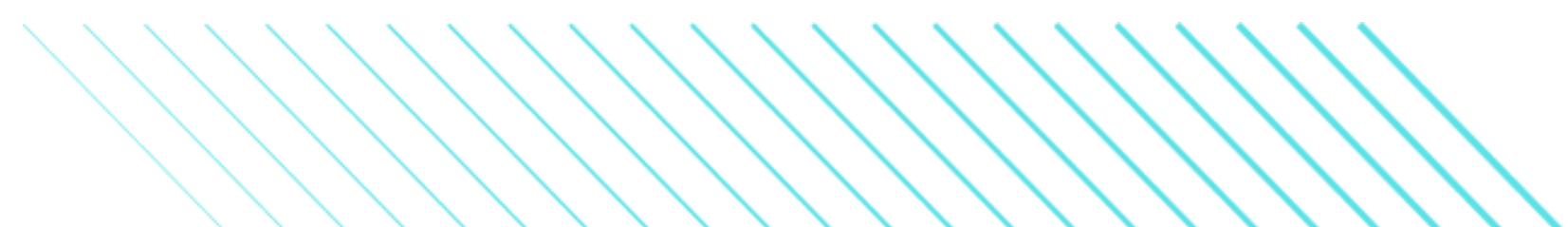
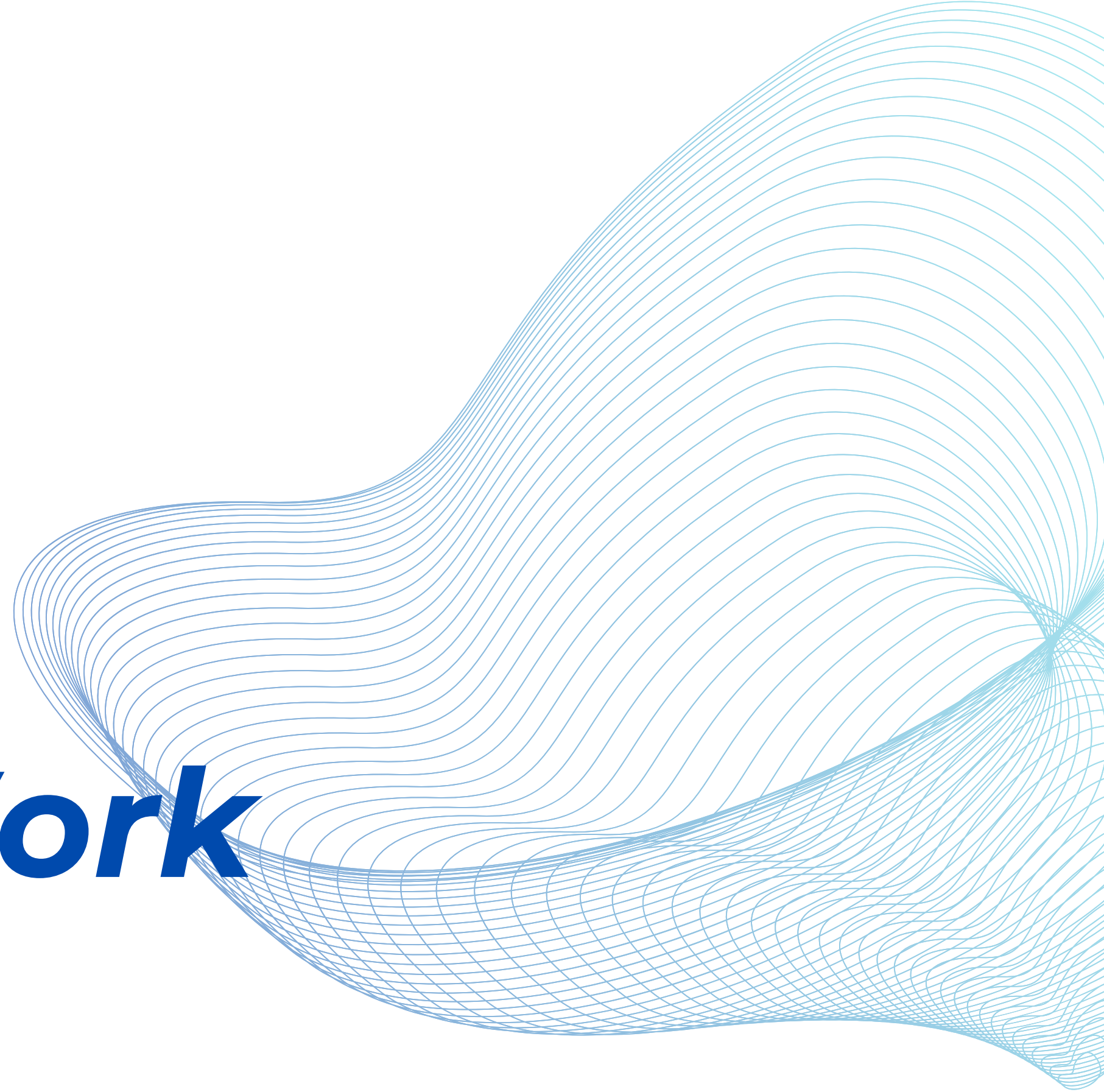
Table 5.4: Coverage diagnostics on different datasets



*ReDial is built from MovieLens dataset

6

Conclusion and Future Work



Conclusions

Answers to research question

RQ1: How can question elicitation promote diverse preference exploration?

Two possible approaches:

- Strategic feature selection (DECREC [11])
 - Broader selection space
 - Feature entropy promote exploration over exploitation
- Intent-dependent multi-path reasoning (CR-Walker [5])
 - Quickly narrow down the search space with KG



Conclusions

Answers to research question

RQ2: How to integrate KG-retrieved info with language model for coherent, grounded CRS?

Two possible approaches:

- Finetune LLM to read and transform KG-retrieved info into meaningful explanation (KECRS [4], KERL [3])
- Utilize retrieved information directly without finetuning (CR-Walker [5], GraphRAG [8])
 - Take advantages of current advancement of LLM



Conclusions

Answers to research question

RQ3: How do retrieval-level and elicitation-level diversity interact for overall diversity?

Elicitation-level diversity **enables and amplifies** retrieval-level diversity through a cascading effect: diverse questions uncover multi-dimensional preferences, which activate multiple retrieval paths in the knowledge graph, leading to comprehensive catalog exploration.



Future work

Direction for thesis

- Focus on **generation-based** approaches for CRS, with attention on **accuracy and coverage on each turn**
- Build a **CRS framework** that separate KG path reasoning retrieval and LM response generation
 - **KG retriever:** utilizing multipath reasoning (CRWalker) and enhance on path selection strategy to adapt to user preference of novelty.
 - **Response generator:**
 - Finetune SLM for specific task: transform KG retrieved information into explainable recommendation
 - OR Utilize existing LLM.
- Deliverables: POC of the framework and evaluation results on existing benchmark



Future work

Direction for thesis

- Focus on **generation-based** approaches for CRS, with attention on **accuracy and coverage on each turn**
- Build a **CRS framework** that separate KG path reasoning retrieval and LM response generation

KG Retriever

utilizing multipath reasoning (CRWalker) and **enhance path selection strategy** to adapt to user preference of novelty

Response generator

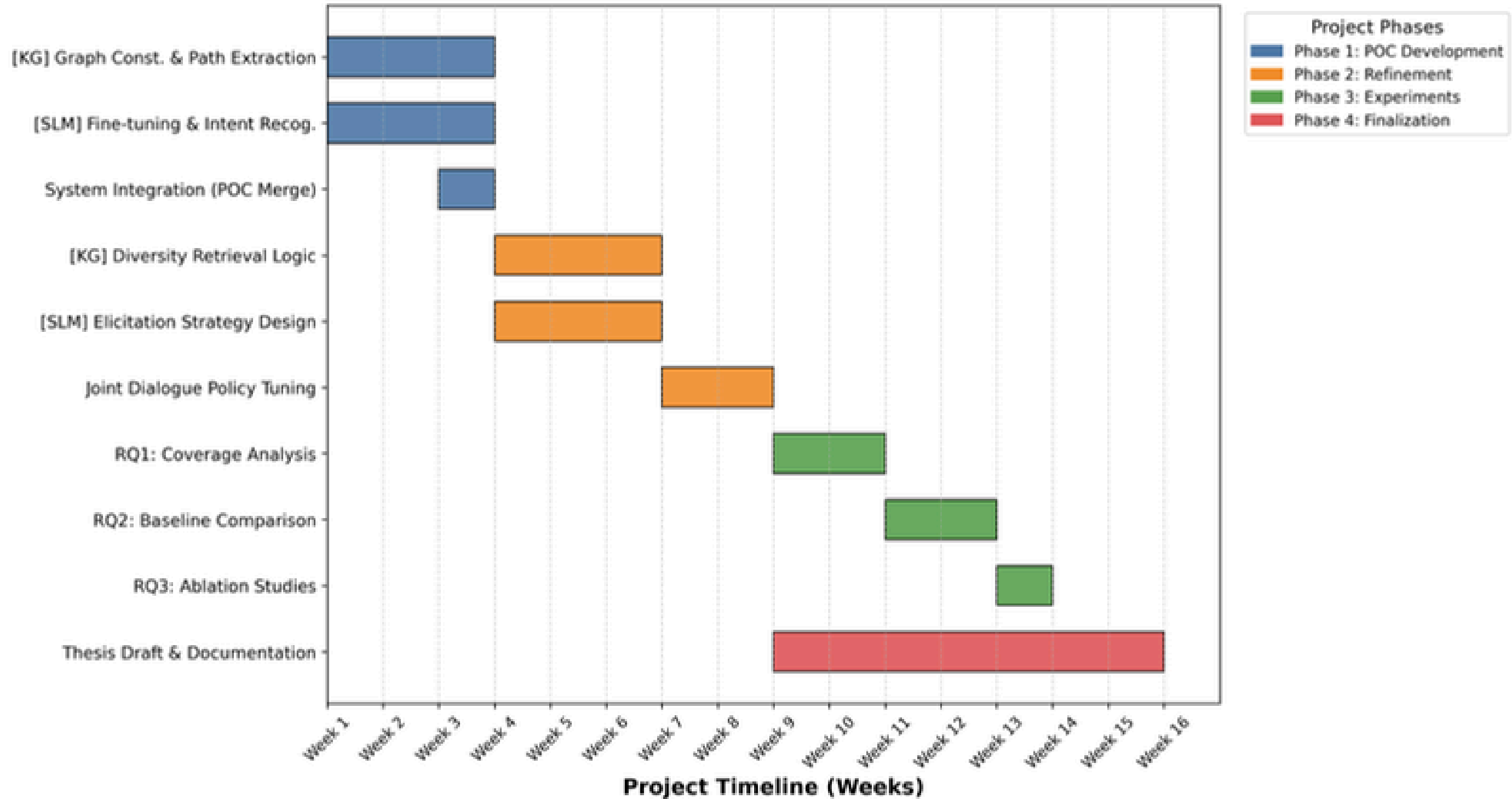
- Finetune SLM for CRS specific context
- OR Utilize existing LLM.

- Deliverables: POC of the framework and evaluation results on existing benchmark



Future plan

Research Implementation Plan (Parallel Workflows)



References

- [1] Q. Chen et al., “Towards knowledge-based recommender dialog system”, in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China: Association for Computational Linguistics, 2019, pp. 1803–1813.
- [2] J. Zhang, Y. Yang, C. Chen, L. He, and Z. Yu, “Kers: A knowledge enhanced framework for recommendation dialog systems with multiple subgoals”, in Findings of the Association for Computational Linguistics: EMNLP2021, PuntaCana, Dominican Republic: Association for Computational Linguistics, 2021, pp. 1092–1101.
- [3] Z. Qiu, L. Luo, and A. W.-C. Liew, “Knowledge graphs and pre-trained language models enhanced representation learning for conversational recommender systems”, IEEE Transactions on Neural Networks and Learning Systems, 2024.
- [4] T. Zhang et al., “Toward knowledge-enriched conversational recommendation system”, in Proceedings of the 4th Workshop on NLP for Conversational AI, Association for Computational Linguistics, 2022, pp. 212-217.
- [5] W. Ma, R. Takanobu, and M. Huang, “Cr-walker: Tree-structured graph reasoning and dialog acts for conversational recommendation”, in Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2021, pp. 1839–1851.



References

- [6] T.Guoetal., “Knowledge graph enhanced language agents for recommen dation”, arXiv preprint arXiv:2410.19627, 2024.
- [7] Z. Qiu, L. Luo, S. Pan, and A. W.-C. Liew, Reasoning over user preferences: Knowledge graph-augmented LLMs for explainable conversational recommendations, 2024. arXiv: 2411.14459 [cs.CL].
- [8] Z. Qiu, L. Luo, Z. Zhao, S. Pan, and A. W.-C. Liew, “Graph retrieval augmented llm for conversational recommender systems”, arXiv preprint arXiv:2503.06430, 2025.
- [9] Y. Jung, E. Jung, and L. Chen, “Towards a unified conversational recom mendation system: Multi-task learning via contextualized knowledge dis tillation”, in Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Singapore: Association for Computational Linguistics, 2023, pp. 13625–13637.
- [10] Y.Deng, Y. Li, F. Sun, B. Ding, and W. Lam, “Unified conversational rec ommendation policy learning via graph-based reinforcement learning”, in Proceedings of the 44th International ACM SIGIR Conference on Re search andDevelopmentinInformation Retrieval (SIGIR '21), ACM,2021, pp. 1431–1441.
- [11] Z. Wang et al., “Diversity-enhanced conversational recommendation via multi-agent reinforcement learning”, Knowledge and Information Systems, vol. 67, no. 9, pp. 7727–7755, 2025.

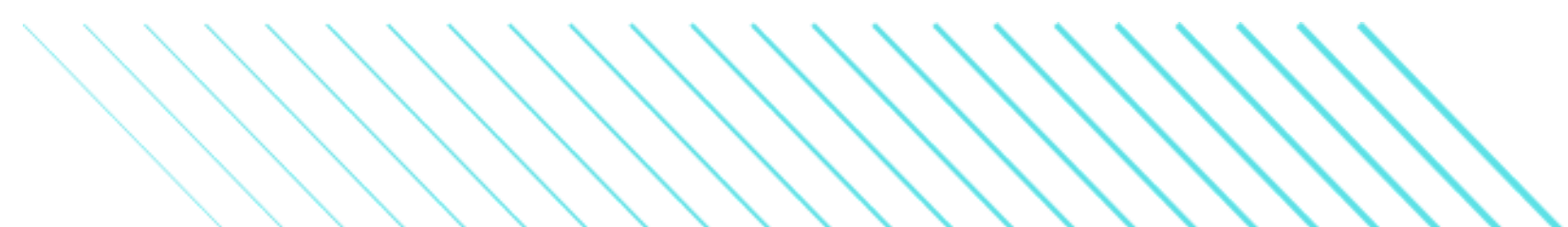
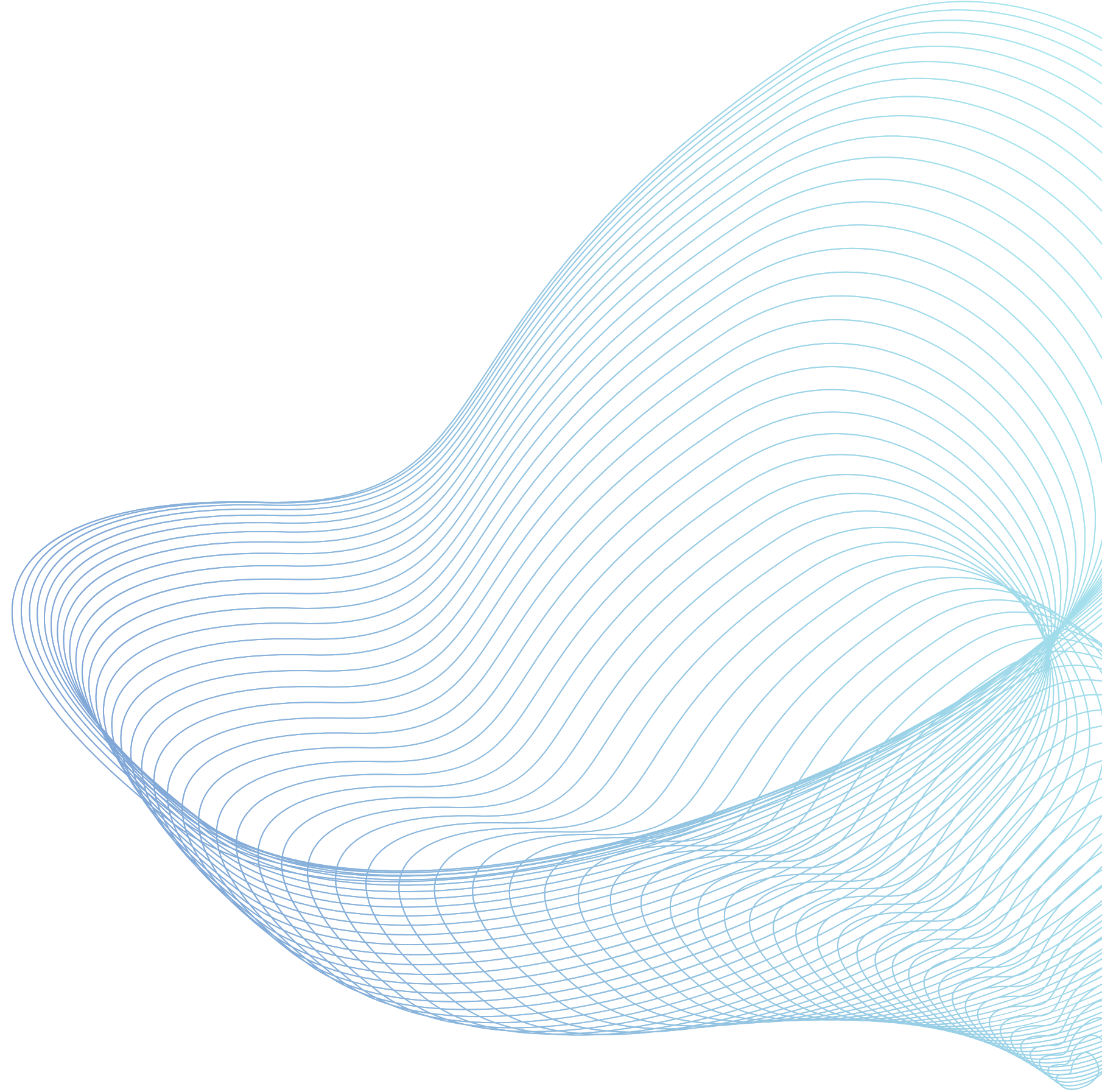


References

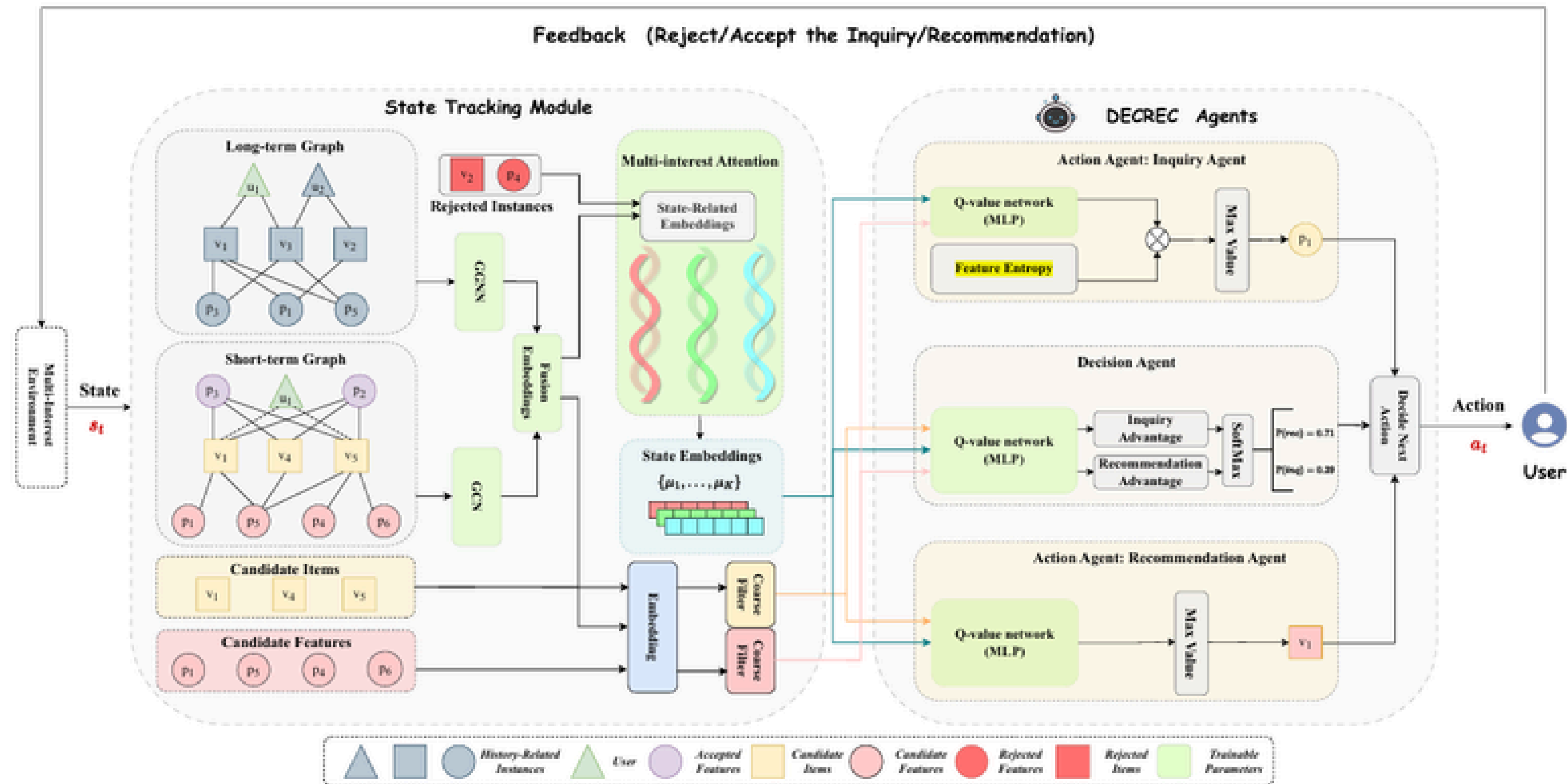
- [12] Y. Xian et al., “Cafe: Coarse-to-fine neural symbolic reasoning for explainable recommendation”, in Proceedings of the 29th ACM International Conference on Information & Knowledge Management(CIKM’20), ACM, 2020, pp. 1645–1654.
- [13] T.-C. Wang, S.-Y. Su, and Y.-N. Chen, “Barcor: Towards a unified framework for conversational recommendation systems”, arXiv preprint arXiv:2203.14257, 2022.
- [14] Z. Zhao, K. Zhou, W. X. Zhao, X. Wang, Y. Cheng, and J.-R. Wen, “Alleviating the long-tail problem in conversational recommender systems”, in Proceedings of the 17th ACM Conference on Recommender Systems, ACM, 2023, pp. 374–385.



***Thank you for
listening***



Diversity-enhanced conversational recommendation via multi-agent reinforcement learning (DECREC)



Language Models

Efficient conversational brain

- Built on **Transformer architecture + attention mechanisms**.
- Evolved from **Large Language Models (LLMs)** to **Small Language Models (SLMs)**.

Lightweight trend enabled by:

- Retrieval-Augmented Generation (RAG)
- Model Context Protocol (MCP) servers

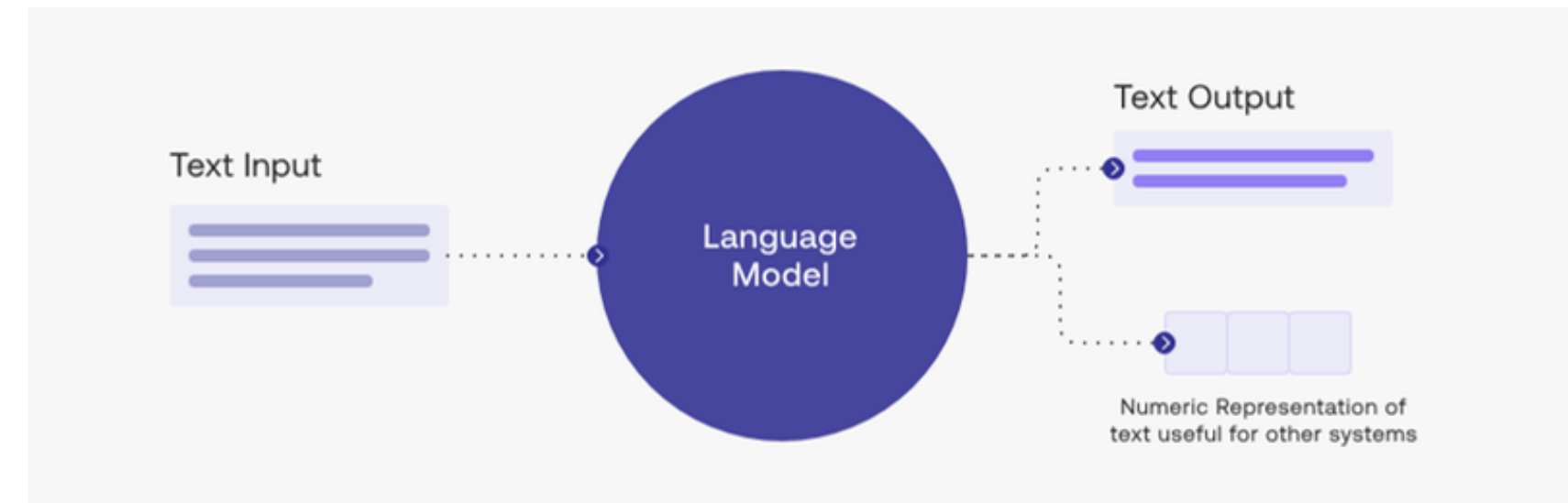


Figure 4: Language Model Overview

Conversation Fluency & Response Quality

- **Metrics:** BLEU, Dist-2, Dist-3, Know-Prec., Know-Rec., Know-F1
- **Paper:** CR-Walker
- **Datasets:** ReDial

Paper / Model	BLEU	Dist-2	Dist-3	Know-Prec.	Know-Rec.	Know-F1
CR-Walker (paper)	26.6	21.2	48.1	46.3	60.3	47.7
CR-Walker (our env)	7.00	16.3	28.9	31.2	40.1	14.2

Table 1: Conversation fluency/quality metrics for CR-Walker on ReDial



Diversity & Coverage metrics

- **Metrics:** I_{num} , I_{suc} , F_{num} , F_{suc} , Cov
- **Paper:** DECREC
- **Datasets:** Amazon-Book, Yelp, LastFM, MovieLens

Paper / Model	Amazon-Book				Yelp			
	I_{num}	I_{suc}	F_{num}	F_{suc}	I_{num}	I_{suc}	F_{num}	F_{suc}
UNICORN (paper)	9789	5972	109	34	16305	6326	53	33
DECREC (paper)	9918	6486	191	45	17459	8648	68	42
DECREC (our env)	10823	8609	775	284	19593	10242	305	128

Table 1: Diversity diagnostics on MRCR benchmarks (Amazon-Book, Yelp)



Diversity & Coverage metrics

- **Metrics:** I_{num} , I_{suc} , F_{num} , F_{suc} , Cov
- **Paper:** DECREC
- **Datasets:** Amazon-Book, Yelp, LastFM, MovieLens

Paper / Model	LastFM				MovieLens			
	I_{num}	I_{suc}	F_{num}	F_{suc}	I_{num}	I_{suc}	F_{num}	F_{suc}
UNICORN (paper)	6249	5211	158	94	2791	2050	27	15
DECREC (paper)	6281	5531	178	99	3701	2786	38	24
DECREC (our env)	7212	6989	705	234	4352	3532	137	62

Table 1: Diversity diagnostics on MCR benchmarks (LastFM, MovieLens)



Scope & Limitations

In-scope

- **Target Task:** Multi-turn Conversational Recommendation (CRS)
 - + Focus: eliciting/clarifying preferences + recommending with brief explanations
- **Methodology Focus:**
 - + SLM-based dialogue: question elicitation + response generation
 - + KG grounding: retrieval-supported facts to reduce hallucination
 - + Diversity-aware design: diversity in elicitation and retrieval
- **Data & Evaluation:**
 - + Offline evaluation on public CRS datasets (e.g., ReDial) + standard CRS metrics

Out-of-scope

- **Online user studies / A-B testing:** no real-user deployment evaluation
- **Very large LLM-centric systems:** focus is on SLM-first designs
- **Perfect factuality & long-context guarantees:** limited by retrieval quality + model capacity

