Large Language Model Powered Agents in the Web

Tutorial at The Web Conference 2024 in Singapore (WWW 2024)

Yang Deng[1], An Zhang[1], Yankai Lin[2], Xu Chen[2], Ji-Rong Wen[2], Tat-Seng Chua[1]

[1] NExT++ Research Centre, National University of Singapore
[2] Gaoling School of Artificial Intelligence, Renmin University of China

dengyang17dydy@gmail.com, an_zhang@nus.edu.sg, yankailin@ruc.edu.cn
xu.chen@ruc.edu.cn, jrwen@ruc.edu.cn, chuats@comp.nus.edu.sg

May 13, 2024, Singapore

Speakers

Homepage

Personal Information

Zhang An 张岸

➢ **Education Background**

- 2021 – present: **Post-Doc**, NUS, School of Computing, NExT++ Research Centre
- 2016 – 2021: **Ph.D**, NUS, Department of Statistics and Data Science
- 2012 – 2016: **B.S.**, Southeast University, School of Mathematics

➢ **Research Interests**: LLM-empowered Agents, Robust and Trustable AI, Recommender System

➢ **Homepage**: https://anzhang314.github.io/

➢ **Email: an_zhang@nus.edu.sg**

Outline

- Part 1: Introduction of LLM-powered Agents
- Part 2: LLM-powered Agents with Tool Learning
- Part 3: LLM-powered Agents in Social Network
- **Part 4: LLM-powered Agents in Recommendation**
- Part 5: LLM-powered Conversational Agents
- Part 6: Open Challenges and Beyond

Significant Gap Between LLMs and Recommender Systems (RecSys)

- **Significant gap between large language models (LLMs) and recommender systems (RecSys).**

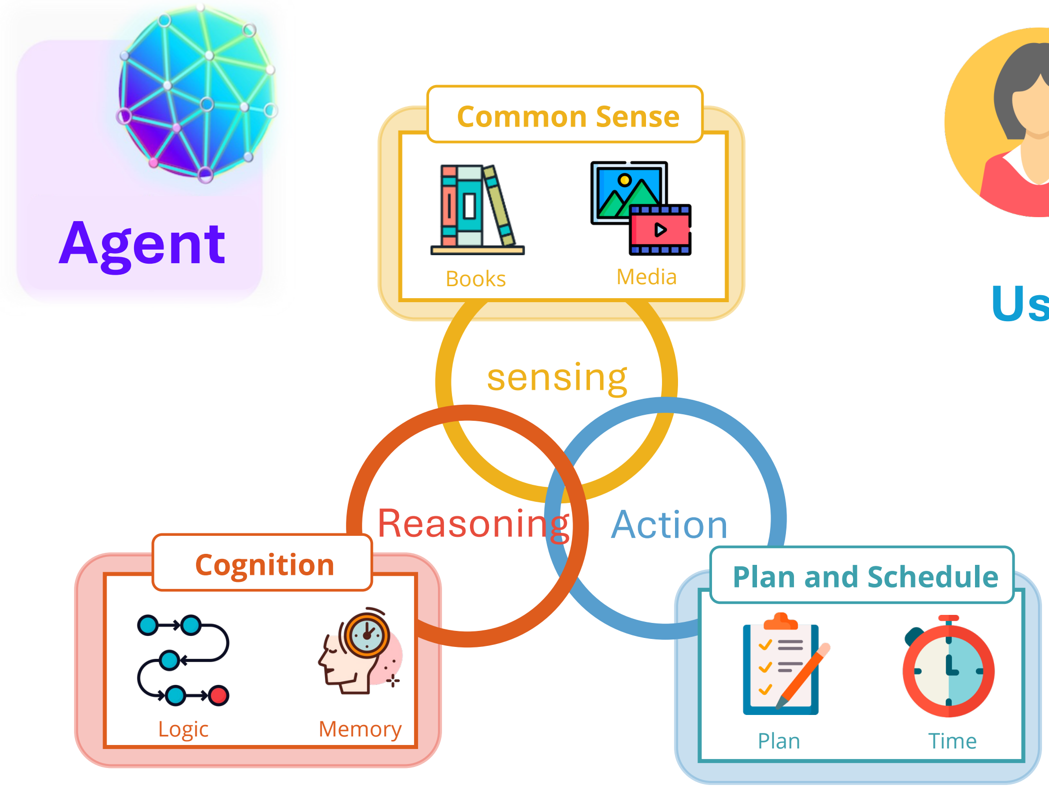
  How to bridge this gap?

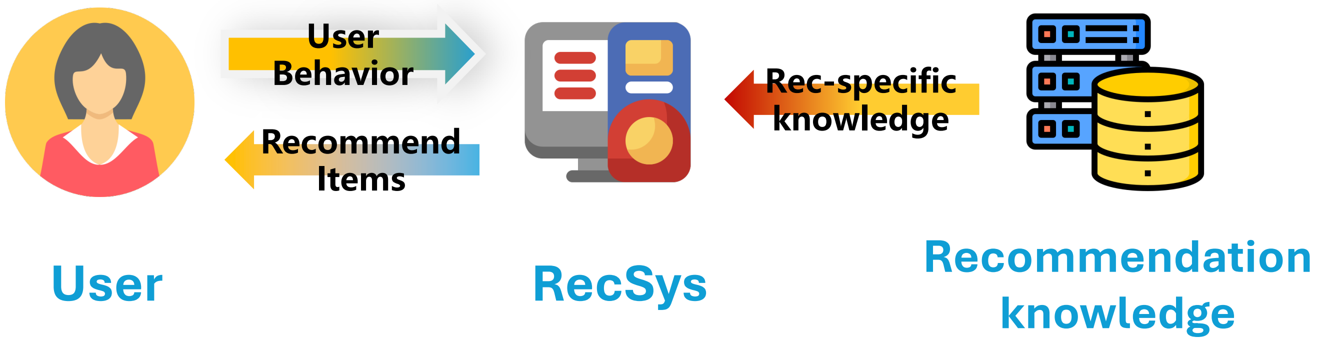
	LLMs	**RecSys**
Scope	Language modelling	User behaviour modelling
Data	Rich world text-based sources	Sparse user-item interactions
Tokens	A chunk of text (Ten thousand level)	Items (Billion level)
Characteristics	General model; Open-world knowledge; High complexity and long inference time;	Leveraging collaborative signals; Lack of cross-domain adaptability; Struggle with cold-start problem; Limited intention understanding;

Significant Gap Between LLMs and Recommender Systems (RecSys)

➢ **Significant gap between large language models (LLMs) and recommender systems (RecSys).**

How to bridge this gap?





- **Two critical components in RecSys:**
  - Understanding **user's behavior/preference**
  - Acquiring recommendation-specific **knowledge**

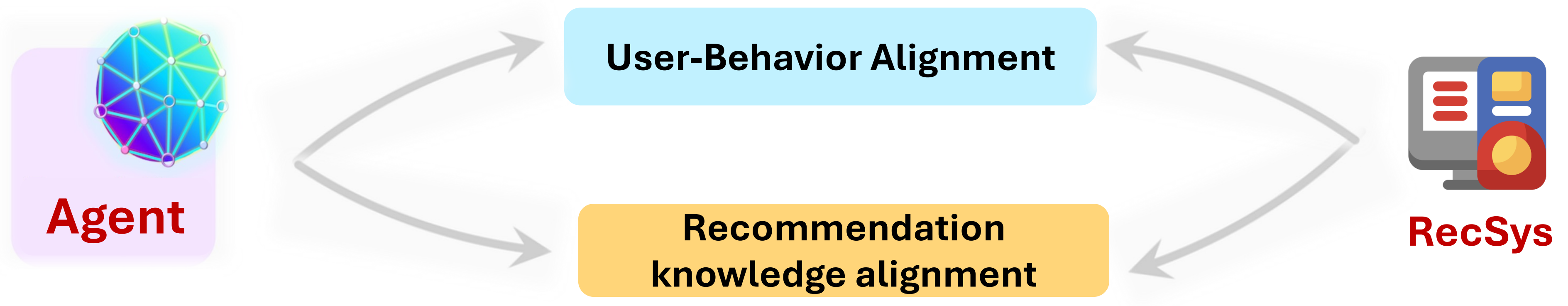
Significant Gap Between LLMs and Recommender Systems (RecSys)

- **Significant gap between large language models (LLMs) and recommender systems (RecSys).**

  How to bridge this gap?

- **Align recommendation space with language space.**
  - **User behavior** alignment
  - **Recommendation knowledge** alignment
- **Two critical components in RecSys:**
  - Understanding **user's behavior/preference**
  - Acquiring recommendation-specific **knowledge**

LLM-powered Agents in Recommendation



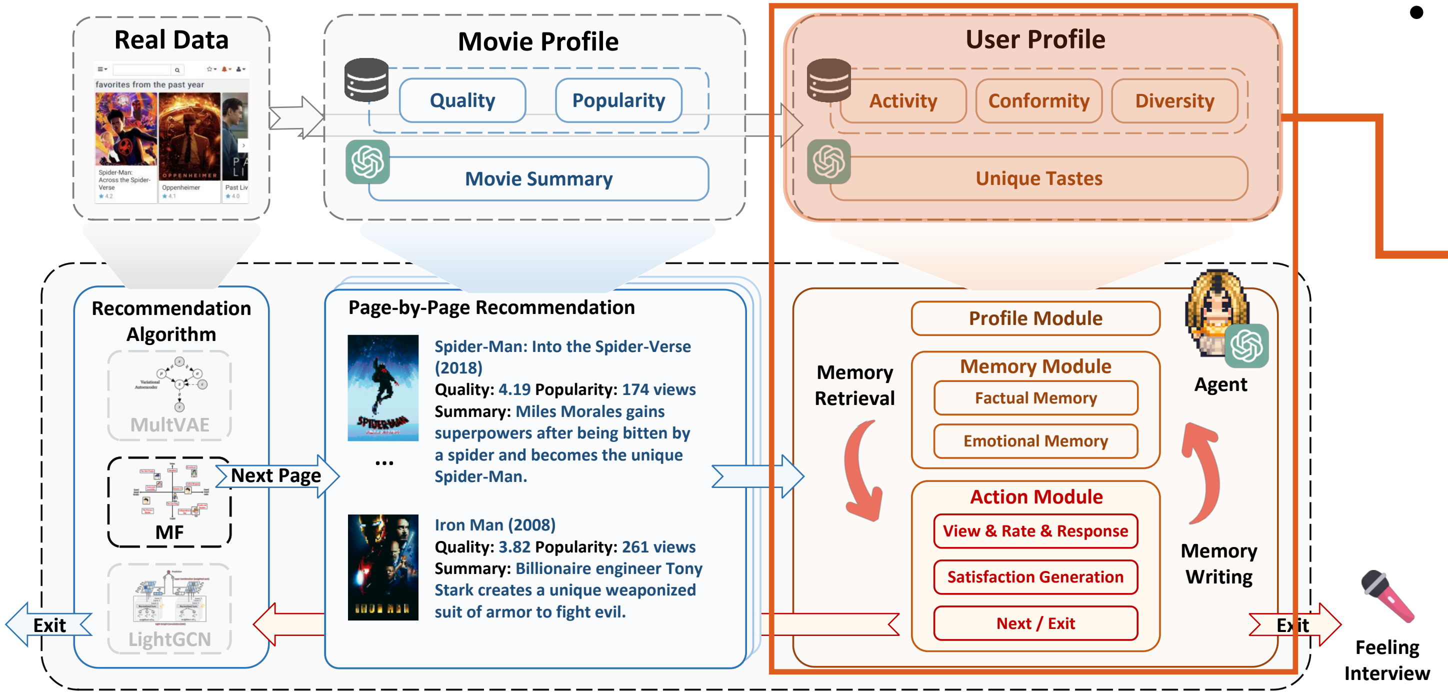
- **LLM-powered Agents have potentials to solve long-standing problems in recommendation**
  - Can an LLM-powered Agent faithfully simulate **users**?
  - Can an LLM-powered Agent be a better **recommender** with recommendation-specific knowledge?

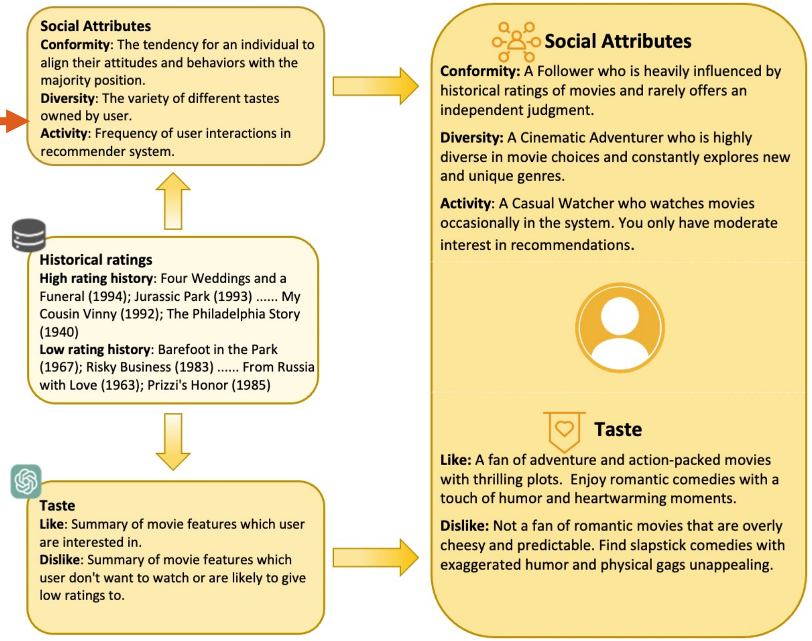
Agents as Users

❑ Agent4Rec: Agent-driven user behavior simulation

- **Key Points:**
  - Can LLM-powered Agent generate faithful user behaviors?

- **User Profile**: **1,000** LLM-empowered generative agents initialized with **real data** in various dataset and augmented by ChatGPT.
- **Item Profile**: Statistical information in dataset and generated summary.





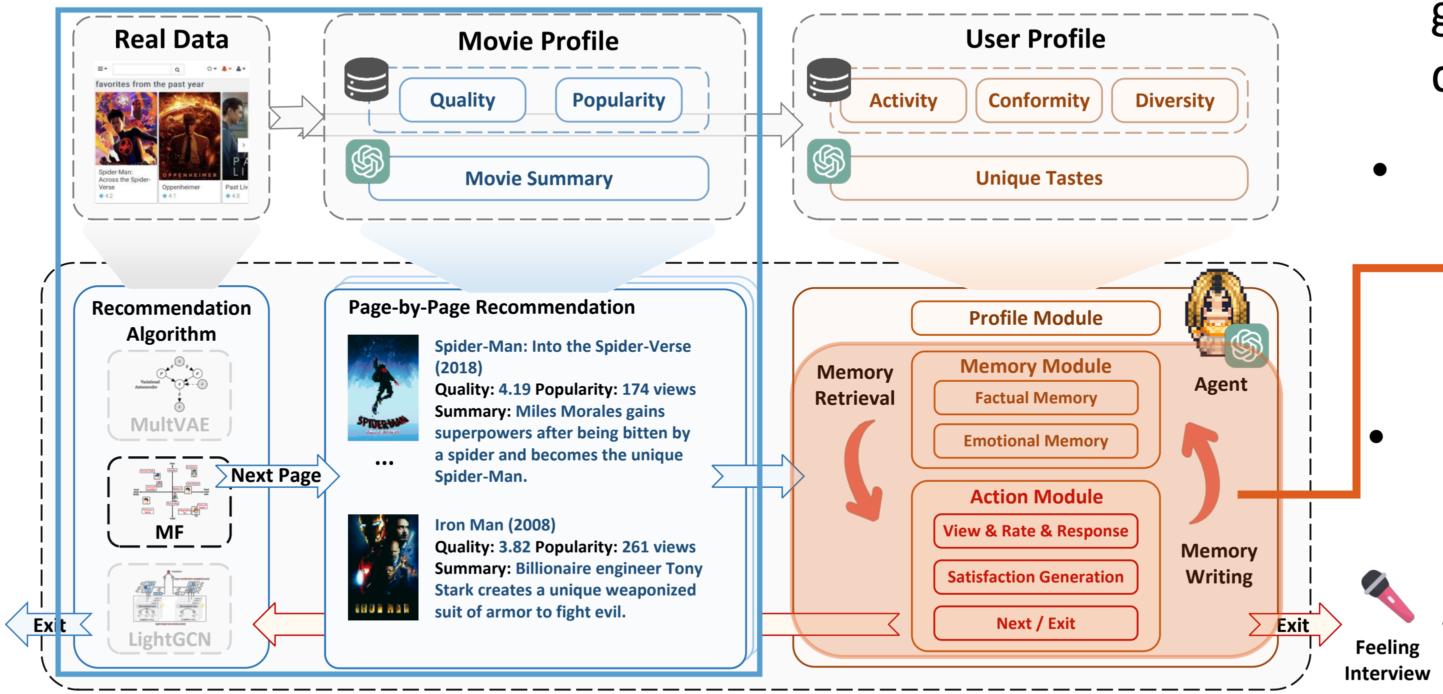
An Zhang et al. On Generative Agents in Recommendation. SIGIR 2024.

Agents as Users
Agent4Rec

Agents as Users

❑ **Agent4Rec: Agent-driven user behavior simulation**

- **Key Points:**
  - Can LLM-powered Agent generate faithful user behaviors?

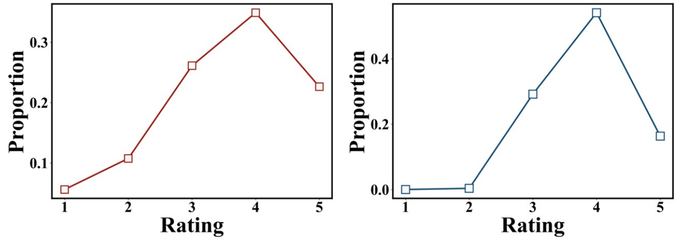


- Agents as users: **1,000** LLM-empowered generative agents initialized from the real dataset.
- **Memory** and **action** modules enable agents to recall past interests and plan future actions (**watch**, **rate**, **evaluate**, **exit**, and **interview**).
- Recommendation environment: Agent4Rec conducts personalized recommendations in a **page-by-page manner** and **pre-implements various recommendation algorithms**.

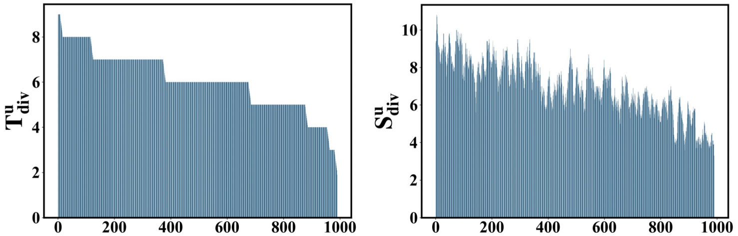
An Zhang et al. On Generative Agents in Recommendation. SIGIR 2024.

Agents as Users
Agent4Rec

- **Key Observations:**
  - Agents are capable of preserving the user's social attributes and preference.
  - Incorporating agents' rating as augmented data can enhance the recommender's performance.



(a) Distribution on MovieLens (b) Agent-simulated distribution

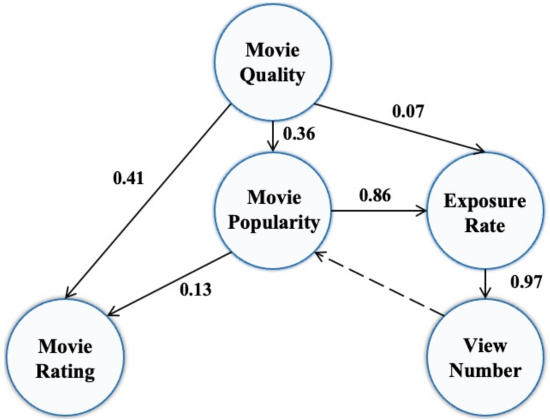


(a) Ground-truth diversity (b) Simulated diversity

Table 3: Page-by-page recommendation enhancement results over various algorithms.

Offline	MF Recall	MF NDCG	MultVAE Recall	MultVAE NDCG	LightGCN Recall	LightGCN NDCG
Origin	0.1506	0.3561	0.1609	0.3512	0.1757	0.3937
+ Viewed	**0.1570***	**0.3604***	**0.1613***	**0.3540***	**0.1765***	**0.3943***
Simulation	$\overline{N}_{exit}$	$\overline{S}_{sat}$	$\overline{N}_{exit}$	$\overline{S}_{sat}$	$\overline{N}_{exit}$	$\overline{S}_{sat}$
Origin	3.17	3.80	3.10	3.75	3.02	3.85
+ Viewed	**3.27***	**3.83***	**3.18***	**3.87***	**3.10***	**3.92***

  - By utilizing ICA-based LiNGAM to analyse the results, we are able to discover **Causal Relations** among movie quality, movie rating, movie popularity, exposure rate, and view number.



- **Offer a simulation platform to test and fine-tune recommender models.**

An Zhang et al. On Generative Agents in Recommendation. SIGIR 2024.

Agents as Users
Agent4Rec

- **Key Observations:**
  - Agents are capable of preserving the user's social attributes and preference.
  - Incorporating agents' rating as augmented data can enhance the recommender's performance.

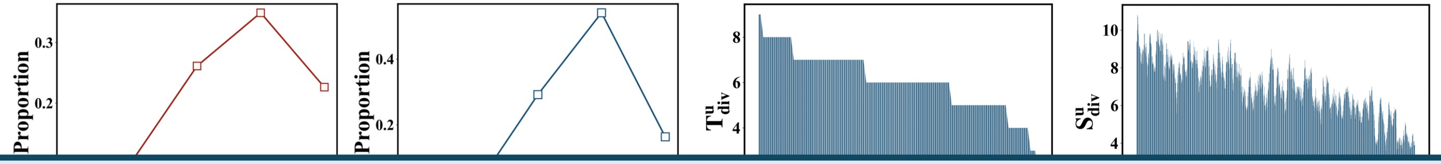
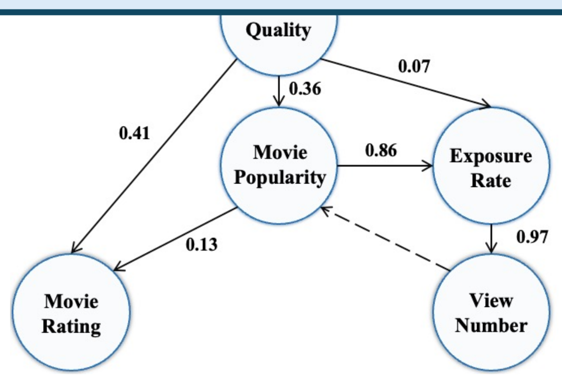


Table 3: Page-by-page recommendation enhancement results over various algorithms.

Offline	MF Recall	MF NDCG	MultVAE Recall	MultVAE NDCG	LightGCN Recall	LightGCN NDCG
Origin	0.1506	0.3561	0.1609	0.3512	0.1757	0.3937

LLM-powered agents are able to generate faithful behaviors.

able to discover **Causal Relations** among movie quality, movie rating, movie popularity, exposure rate, and view number.



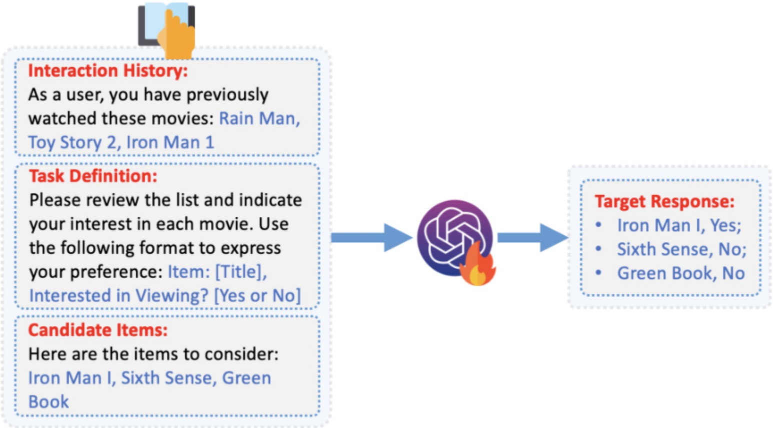
- **Offer a simulation platform to test and fine-tune recommender models.**

An Zhang et al. On Generative Agents in Recommendation. SIGIR 2024.

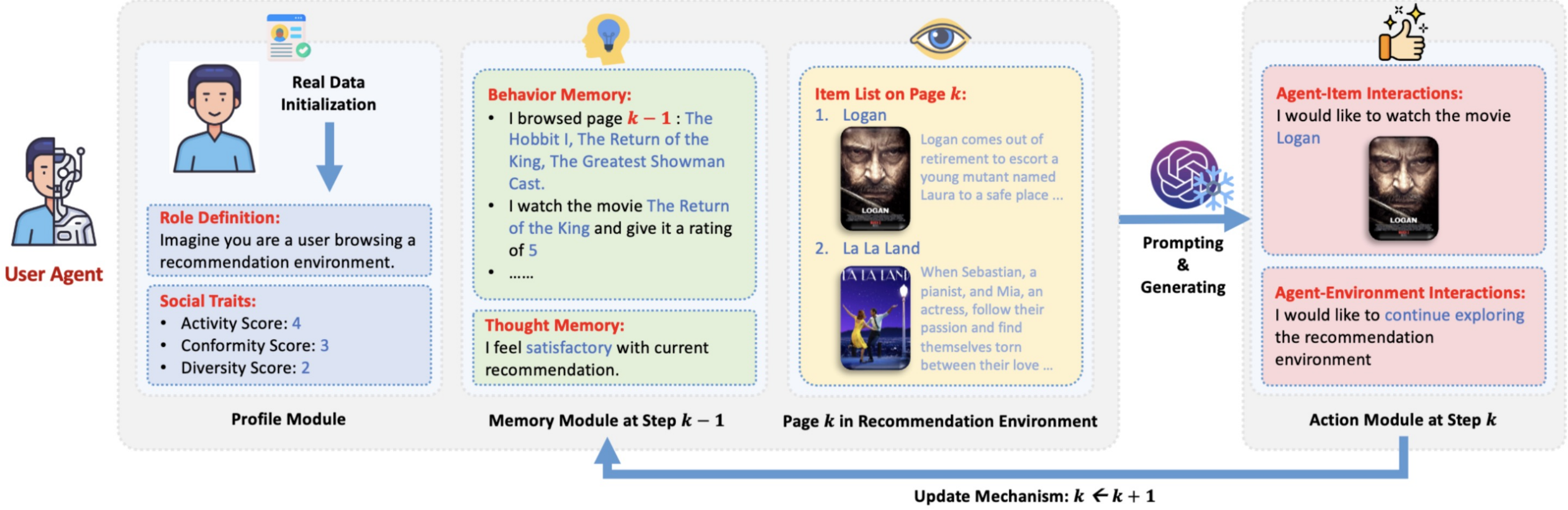
Agents as Users
UGen

Agents as Users

- **Key Points :**
  - Can LLM-powered Agents generated behaviors benefit the recommender?
  - Cooperating updated Agent4Rec framework with finetuning GPT-3.5-turbo as a warmup, agents can accurately select their interested items among candidate set.



- Agents have potentials to replace discriminative learning with generative learning paradigms for user modeling in recommendation.
- Conduct extensive experiments on three dataset from different domains (movie, book, game).



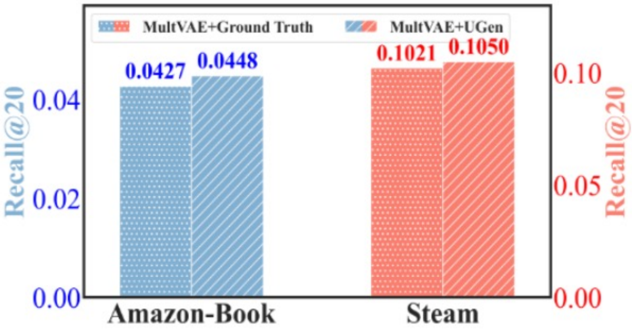
Agents as Users
UGen

- **Key Observations:**
  - Agents are capable of providing effective behaviors, especially in scenarios with sparse data.

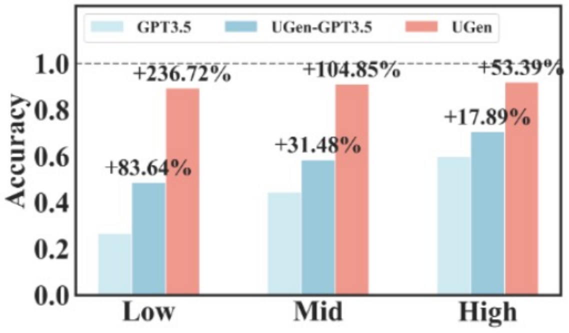
Table 2: Faithfulness Evaluation of Agent's Behavior Alignment with Real User Preferences. Average ground-truth positives are 7.14 (MovieLens), 6.57 (Amazon-Book), and 5.80 (Steam). UGen shows significant improvement with p-value << 0.05.

	MovieLens				Amazon-Book				Steam			
	Acc	Pre	Rec	#Select	Acc	Pre	Rec	#Select	Acc	Pre	Rec	#Select
GPT3.5	0.5295	0.4307	0.7369	11.63	0.4202	0.3855	**0.9072**	17.10	0.4350	0.3430	0.9164	16.59
GPT4	0.6930	0.5743	0.6577	7.00	0.7947	0.6500	0.6003	5.16	0.7844	0.5103	0.7072	6.22
RecAgent	0.6168	0.4519	**0.8921**	13.95	0.5411	0.3714	0.8150	14.65	0.4916	0.3485	**0.9389**	15.55
RAH	0.5758	0.4096	0.6383	9.44	0.7253	0.3355	0.3950	7.45	0.6118	0.3874	0.6262	10.37
UGen-GPT3.5	0.7002	0.4999	0.8600	12.02	0.5690	0.3989	0.8771	14.52	0.5308	0.3688	0.9387	14.74
UGen-GPT4	0.8030	0.5903	0.8142	8.14	0.8419	0.6539	0.7894	8.49	0.8210	0.5306	0.8210	8.85
UGen-Gemini	0.7556	0.4643	0.5021	7.44	0.8375	0.6562	0.6086	4.00	0.7650	0.5286	0.6940	8.80
UGen	**0.9255**	**0.8004**	0.5352	4.55	**0.9171**	**0.7579**	0.6667	5.71	**0.9009**	**0.7007**	0.6895	5.54

	MovieLens		Amazon-Book		Steam	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
MF	0.1529	0.3186	0.0257	0.0480	0.0694	0.0567
+ Random	0.1365	0.2913	0.0199	0.0225	0.0526	0.0432
+ GPT3.5	0.1448	0.3089	0.0253	0.0330	0.0732	0.0608
+ RecAgent	0.1400	0.2990	0.0254	0.0317	0.0696	0.0567
+ RAH	0.1363	0.2917	0.0257	0.0370	0.0731	0.0604
+ UGen	**0.1667**	**0.3396**	**0.0413**	**0.0573**	**0.0807**	**0.0659**
Imp.% over MF	*9.03%*	*6.59%*	*60.70%*	*19.38%*	*16.28%*	*16.23%*
MultVAE	0.1668	0.3107	0.0342	0.0559	0.0816	0.0666
+ Random	0.1630	0.3027	0.0226	0.0218	0.0752	0.0581
+ GPT3.5	0.1708	0.3188	0.0329	0.0336	0.0878	0.0717
+ RecAgent	0.1723	0.3202	0.0292	0.0403	0.0883	0.0716
+ RAH	0.1693	0.3183	0.0320	0.0388	0.0939	0.0774
+ UGen	**0.1725**	**0.3202**	**0.0448**	**0.0612**	**0.1050**	**0.0854**
Imp.% over MultVAE	*2.15%*	*3.06%*	*30.99%*	*9.48%*	*28.68%*	*28.23%*
LightGCN	0.1847	0.3628	0.0420	0.0670	0.0886	0.0757
+ Random	0.1650	0.3358	0.0257	0.0354	0.0762	0.0604
+ GPT3.5	0.1693	0.3462	0.0408	0.0536	0.0817	0.0694
+ RecAgent	0.1650	0.3393	0.0386	0.0518	0.0802	0.0668
+ RAH	0.1597	0.3340	0.0391	0.0542	0.0867	0.0719
+ UGen	**0.1899**	**0.3722**	**0.0555**	**0.0752**	**0.1140**	**0.0952**
Imp.% over LightGCN	*2.82%*	*2.59%*	*32.14%*	*12.24%*	*28.67%*	*25.76%*



(a) Augmented MultVAE



(b) Accuracy on Amazon-Book

Table 4: Human Evaluation on Steam

	Random	Pop	MF	MF+Full	MF+Human
Average Rank	4.72	3.22	2.61	2.50	**1.94**

Agents as Users
UGen

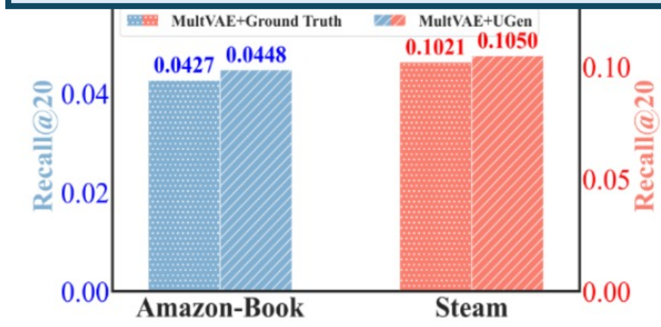
- **Key Observations:**
  - Agents are capable of providing effective behaviors, especially in scenarios with sparse data.

Table 2: Faithfulness Evaluation of Agent's Behavior Alignment with Real User Preferences. Average ground-truth positives are 7.14 (MovieLens), 6.57 (Amazon-Book), and 5.80 (Steam). UGen shows significant improvement with p-value << 0.05.

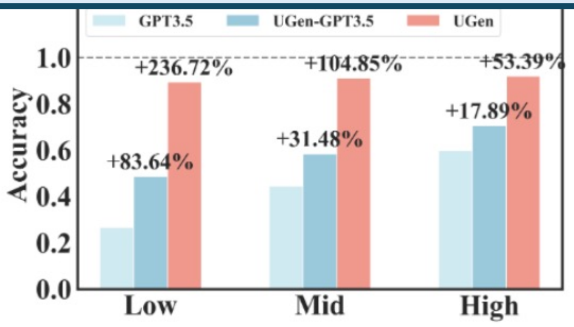
	MovieLens				Amazon-Book				Steam			
	Acc	Pre	Rec	#Select	Acc	Pre	Rec	#Select	Acc	Pre	Rec	#Select
GPT3.5	0.5295	0.4307	0.7369	11.63	0.4202	0.3855	**0.9072**	17.10	0.4350	0.3430	0.9164	16.59
GPT4	0.6930	0.5743	0.6577	7.00	0.7947	0.6500	0.6003	5.16	0.7844	0.5103	0.7072	6.22
RecAgent	0.6168	0.4519	**0.8921**	13.95	0.5411	0.3714	0.8150	14.65	0.4916	0.3485	**0.9389**	15.55
RAH	0.5758	0.4096	0.6383	9.44	0.7253	0.3355	0.3950	7.45	0.6118	0.3874	0.6262	10.37

	MovieLens		Amazon-Book		Steam	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
MF	0.1529	0.3186	0.0257	0.0480	0.0694	0.0567
+ Random	0.1365	0.2913	0.0199	0.0225	0.0526	0.0432
+ GPT3.5	0.1448	0.3089	0.0253	0.0330	0.0732	0.0608
+ RecAgent	0.1400	0.2990	0.0254	0.0317	0.0696	0.0567
+ RAH	0.1363	0.2917	0.0257	0.0370	0.0731	0.0604
+ UGen	**0.1667**	**0.3396**	**0.0413**	**0.0573**	**0.0807**	**0.0659**
Imp.% over MF	*9.03%*	*6.59%*	*60.70%*	*19.38%*	*16.28%*	*16.23%*
+ Random	0.1650	0.3358	0.0257	0.0354	0.0762	0.0604
+ GPT3.5	0.1693	0.3462	0.0408	0.0536	0.0817	0.0694
+ RecAgent	0.1650	0.3393	0.0386	0.0518	0.0802	0.0668
+ RAH	0.1597	0.3340	0.0391	0.0542	0.0867	0.0719
+ UGen	**0.1899**	**0.3722**	**0.0555**	**0.0752**	**0.1140**	**0.0952**
Imp.% over LightGCN	*2.82%*	*2.59%*	*32.14%*	*12.24%*	*28.67%*	*25.76%*

Behaviors generated by LLM-powered agents can benefit recommenders.



(a) Augmented MultVAE



(b) Accuracy on Amazon-Book

Table 4: Human Evaluation on Steam

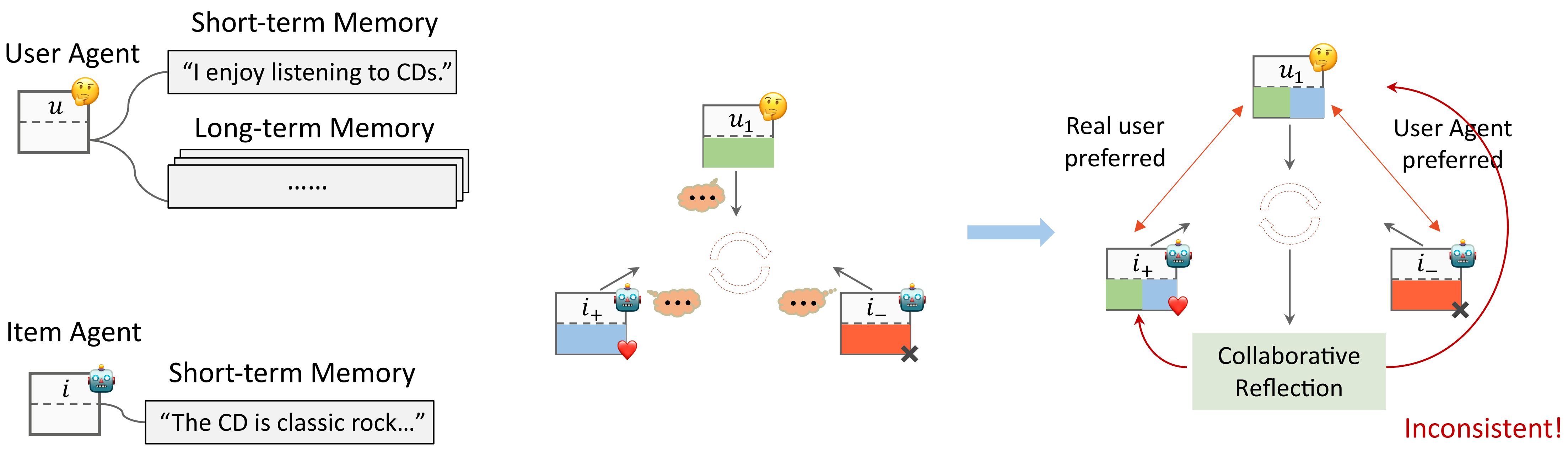
	Random	Pop	MF	MF+Full	MF+Human
Average Rank	4.72	3.22	2.61	2.50	**1.94**

Agents as Users & Items
AgentCF

Agents as Users & Items

❑ **AgentCF: text-based collaborative learning**

- **Key Points:**
  - Can LLM-powered Agent simulate collaborative signals/user-item interactions?



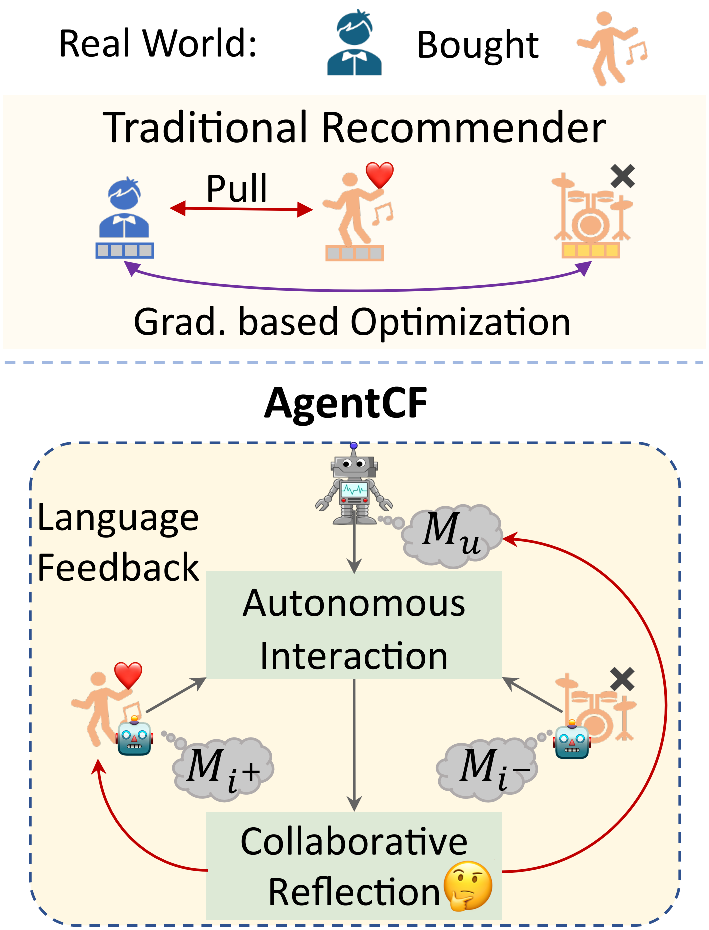
Junjie Zhang et al. AgentCF: Collaborative Learning with Autonomous Language Agents for Recommender Systems. WWW 2024.

Agents as Users & Items
AgentCF

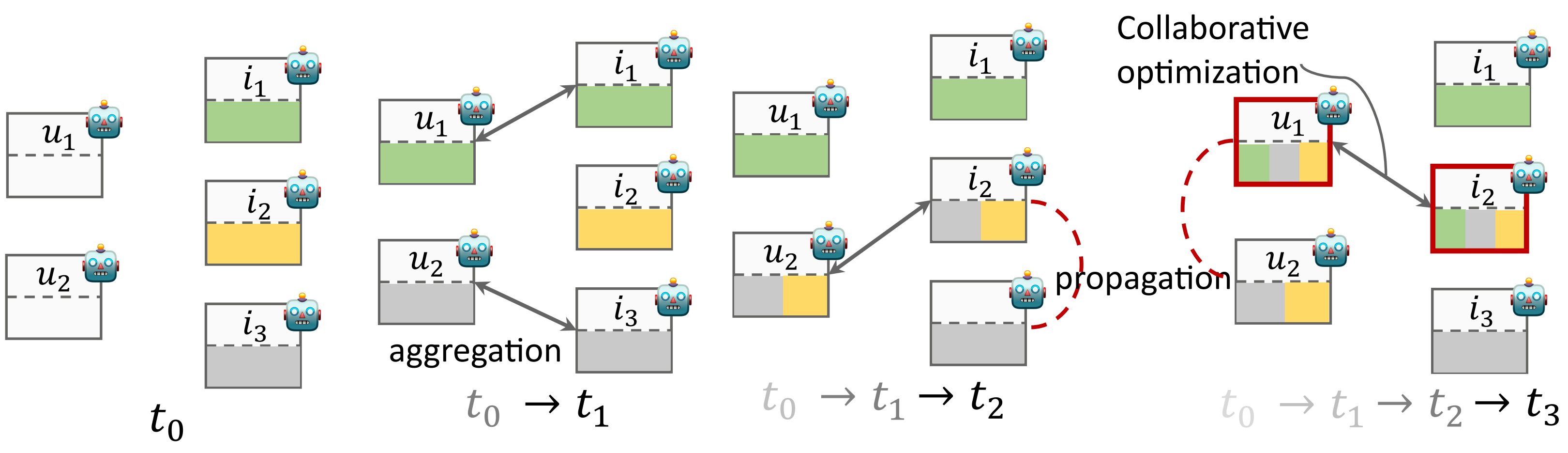
Agents as Users & Items

❑ **AgentCF: text-based collaborative learning**

- **Key Points:**
  - Can LLM-powered Agent simulate collaborative signals/user-item interactions?



- **Key idea**: Parameter-free text-based collaborative optimization.



Junjie Zhang et al. AgentCF: Collaborative Learning with Autonomous Language Agents for Recommender Systems. WWW 2024.

Agents as Users & Items
AgentCF

- **Key Observations:**
  - Agents are capable of simulating user-item interactions.

Method	CDs_{sparse}			CDs_{dense}			$Office_{sparse}$			$Office_{dense}$		
	N@1	N@5	N@10	N@1	N@5	N@10	N@1	N@5	N@10	N@1	N@5	N@10
BPR_{full}	0.1900	0.4902	0.5619	0.3900	0.6784	0.7089	0.1600	0.3548	0.4983	0.5600	0.7218	0.7625
$SASRec_{full}$	0.3300	0.5680	0.6381	0.5800	0.7618	0.7925	0.2500	0.4106	0.5467	0.4700	0.6226	0.6959
BPR_{sample}	0.1300	0.3597	0.4907	0.1300	0.3485	0.4812	0.0100	0.2709	0.4118	0.1200	0.2705	0.4576
$SASRec_{sample}$	0.1900	0.3948	0.5308	0.1300	0.3151	0.4676	0.0700	0.2775	0.4437	**0.3600**	**0.5027**	**0.6137**
Pop	0.1100	0.2802	0.4562	0.0400	0.1504	0.3743	0.1100	0.2553	0.4413	0.0700	0.2273	0.4137
BM25	0.0800	0.3066	0.4584	0.0600	0.2624	0.4325	0.1200	0.2915	0.4693	0.0600	0.3357	0.4540
LLMRank	0.1367	0.3109	0.4715	0.1333	0.3689	0.4946	0.1750	0.3340	0.4728	0.2067	0.3881	0.4928
$AgentCF_{B}$	0.1900	0.3466	0.5019	0.2067	0.4078	0.5328	0.1650	0.3359	0.4781	0.2067	0.4217	0.5335
$AgentCF_{B+R}$	**0.2300**	**0.4373**	**0.5403**	**0.2333**	0.4142	**0.5405**	0.1900	0.3589	0.5062	0.1933	0.3916	0.5247
$AgentCF_{B+H}$	0.1500	0.4004	0.5115	0.2100	**0.4164**	0.5198	**0.2133**	**0.4379**	**0.5076**	0.1600	0.3986	0.5147

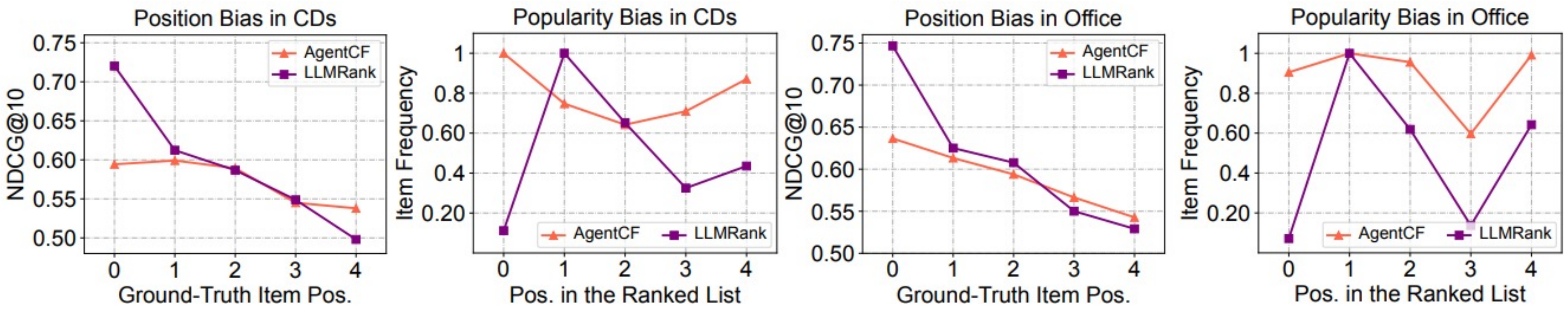


Figure 2: Analysis of whether our approach can simulate personalized agents to mitigate position bias and popularity bias.

Agents as Users & Items
AgentCF

- **Key Observations:**
  - Agents are capable of simulating user-item interactions.

Method	CDs$_{sparse}$			CDs$_{dense}$			Office$_{sparse}$			Office$_{dense}$		
	N@1	N@5	N@10	N@1	N@5	N@10	N@1	N@5	N@10	N@1	N@5	N@10
BPR$_{full}$	0.1900	0.4902	0.5619	0.3900	0.6784	0.7089	0.1600	0.3548	0.4983	0.5600	0.7218	0.7625
SASRec$_{full}$	0.3300	0.5680	0.6381	0.5800	0.7618	0.7925	0.2500	0.4106	0.5467	0.4700	0.6226	0.6959
BPR$_{sample}$	0.1300	0.3597	0.4907	0.1300	0.3485	0.4812	0.0100	0.2709	0.4118	0.1200	0.2705	0.4576
SASRec$_{sample}$	0.1900	0.3948	0.5308	0.1300	0.3151	0.4676	0.0700	0.2775	0.4437	**0.3600**	**0.5027**	**0.6137**

Agents can faithfully simulate user-item interactions.

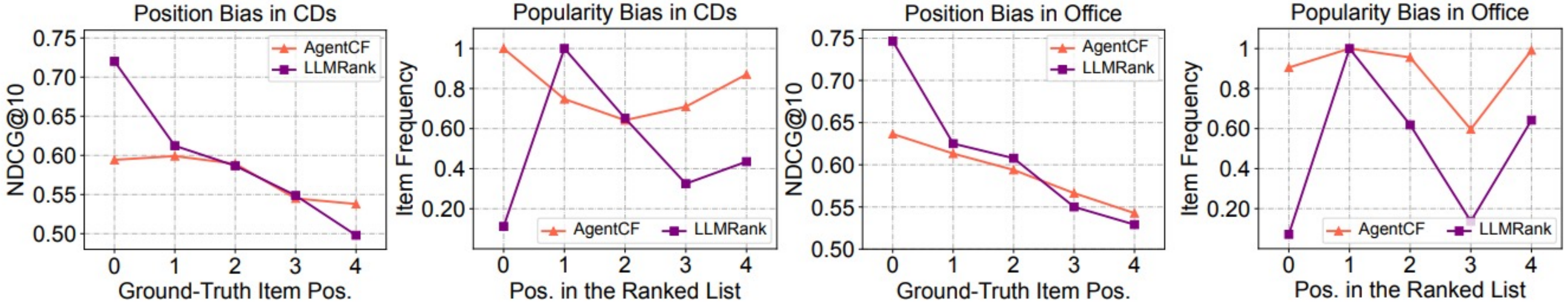
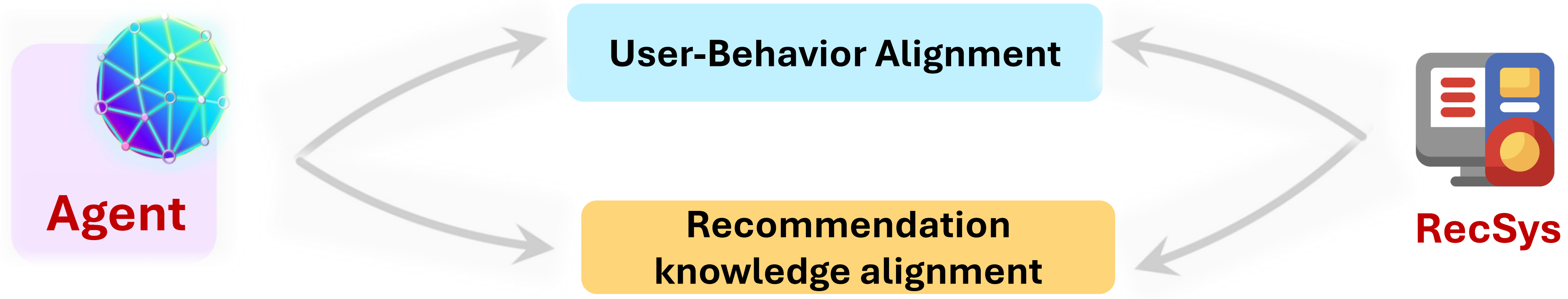


Figure 2: Analysis of whether our approach can simulate personalized agents to mitigate position bias and popularity bias.

LLM-powered Agents in Recommendation



- **LLM-empowered have potentials to solve long-standing problems in recommendation**
  - Can an LLM-powered Agent faithfully simulate **users**?
    - **Agent4Rec, UGen, AgentCF, RecAgent**
  - Can an LLM-powered Agent be a better **recommender** with recommendation-specific knowledge?

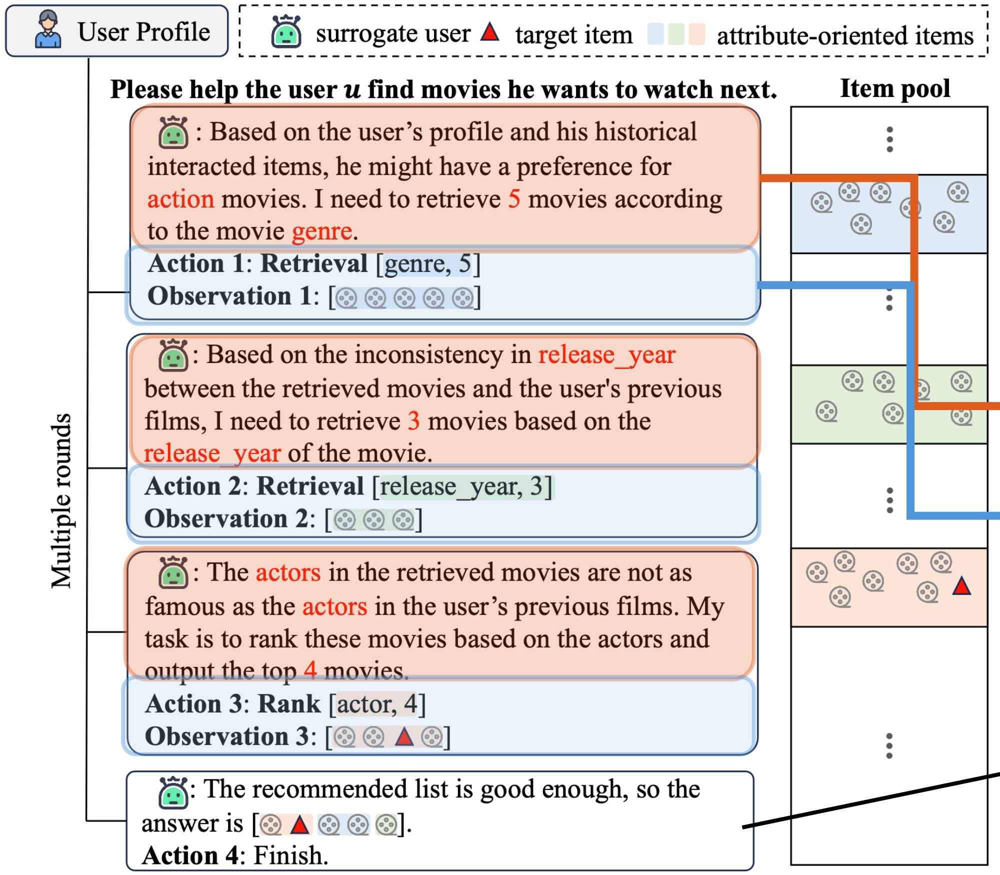
Agent as Recommender
ToolRec

Agent as Recommender

❑ **ToolRec: Tool-enhanced LLM-based recommender**

▪ **Key Points:**

- Can Agents Utilize External Tools to Enhance Recommendations?



Key Idea:

- **Use LLMs to understand current contexts and preferences, and apply attribute-oriented tools to find suitable items.**

Two stages:

- Learning Preferences: LLM-based surrogate user learns user preferences and makes decisions
- Exploration of Items: uses attribute-oriented tools to explore a wide range of items

❖ Process finishes when the LLM-based surrogate user is satisfied with the item list

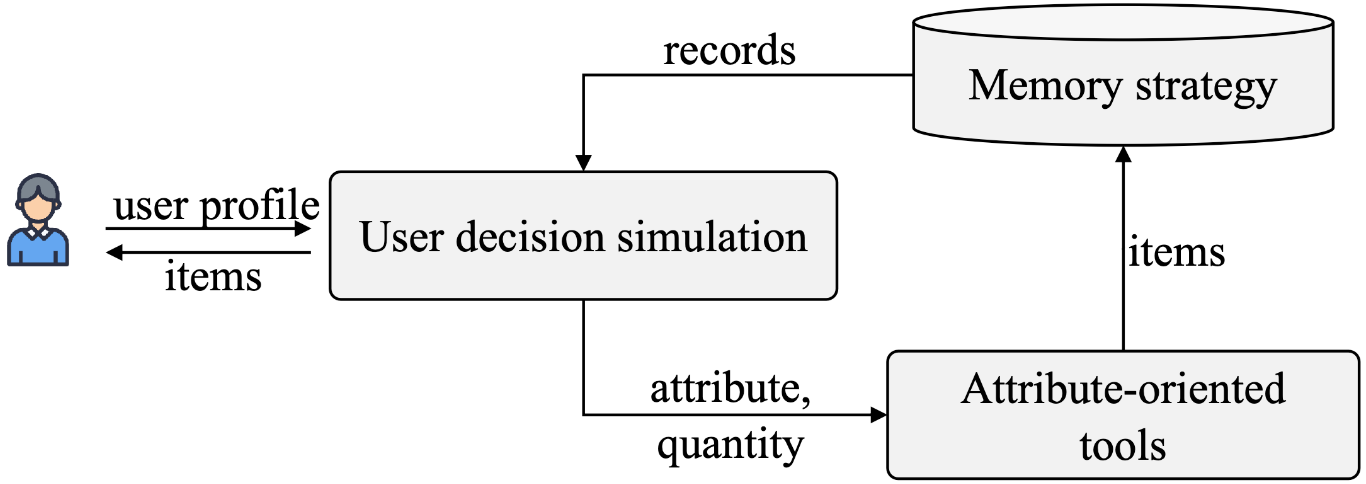
Yuyue Zhao et al. Let Me Do It For You: Towards LLM Empowered Recommendation via Tool Learning. SIGIR 2024.

Agent as Recommender
ToolRec

Agent as Recommender

❑ **ToolRec: Tool-enhanced LLM-based recommender**

- **Key Points:**
  - Can Agents Utilize External Tools to Enhance Recommendations?



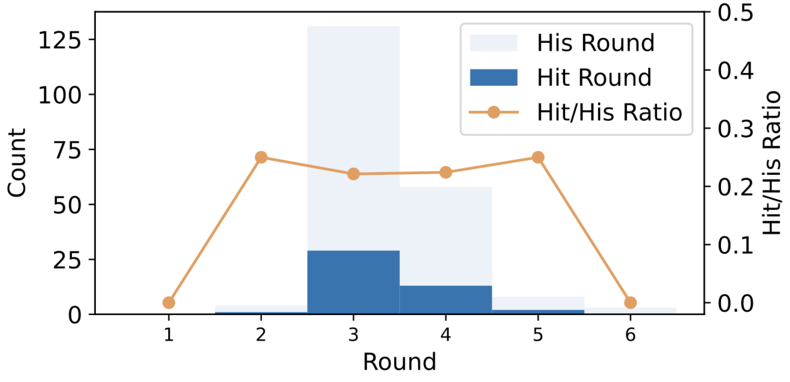
- LLMs as the central controller, simulating the user decision.
- Attribute-oriented Tools: rank tools & retrieval tools.
- Memory strategy can ensure the correctness of generated items and cataloging candidate items.

Yuyue Zhao et al. Let Me Do It For You: Towards LLM Empowered Recommendation via Tool Learning. SIGIR 2024.

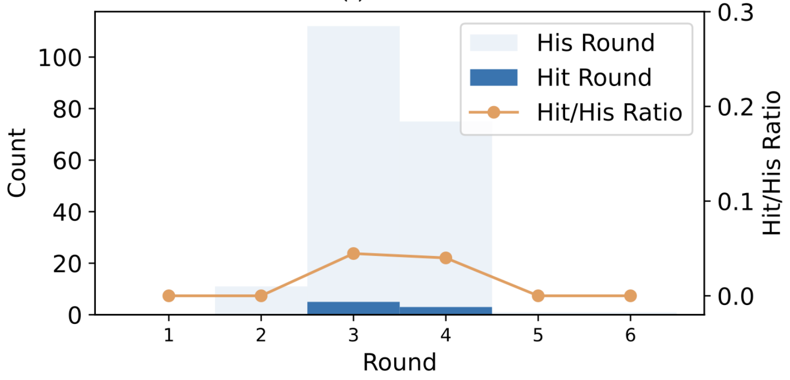
Agent as Recommender
ToolRec

- **Key Observations:**
  - Benefiting from rank tools and tools, ToolRec excels on the ML-1M and Amazon-Book datasets compared to baseline recommenders, demonstrating that it can better align with the users' intent.

	ML-1M		Amazon-Book		Yelp2018	
	Recall	NDCG	Recall	NDCG	Recall	NDCG
SASRec	0.203±0.047	<u>0.1017</u>±0.016	0.047±0.015	0.0205±0.006	0.030±0.005	0.0165±0.006
BERT4Rec	0.158±0.024	0.0729±0.008	0.042±0.015	0.0212±0.009	<u>0.033</u>±0.021	**0.0218**±0.016
P5	<u>0.208</u>±0.021	0.0962±0.009	0.006±0.003	0.0026±0.002	0.012±0.005	0.005±0.001
SASRec_{BERT}	0.192±0.015	0.0967±0.006	0.042±0.003	0.0194±0.002	0.032±0.016	0.0131±0.007
BERT4Rec_{BERT}	0.202±0.013	0.0961±0.009	0.045±0.023	0.0233±0.012	**0.040**±0.028	<u>0.0208</u>±0.015
Chat-REC	0.185±0.044	0.1012±0.016	0.033±0.015	0.0171±0.007	0.022±0.003	0.0121±0.001
LLMRank	0.183±0.049	0.0991±0.020	<u>0.047</u>±0.013	<u>0.0246</u>±0.004	0.030±0.005	0.0140±0.004
ToolRec	**0.215**±0.044	**0.1171**±0.018	**0.053**±0.013	**0.0259**±0.005	0.028±0.003	0.0159±0.001
ToolRec_{B}	0.185±0.018	0.0895±0.002	0.043±0.013	0.0223±0.008	0.025±0.005	0.0136±0.009
Improvement	3.36%	15.10%	14.28%	5.14%	−29.16%	−27.32%



(a) ML-1M.



(b) Amazon-Book.

  - ToolRec shows subpar performance on the Yelp2018 dataset - local (niche) businesses.
  - Most processes conclude in three or four rounds, indicating that the LLM can understand user preferences after a few iterations.

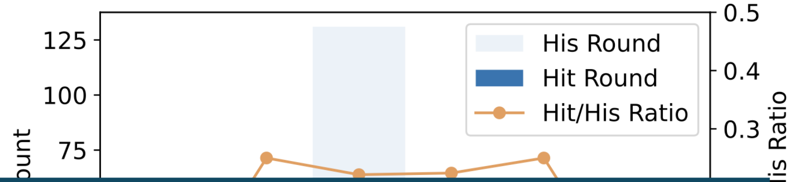
Yuyue Zhao et al. Let Me Do It For You: Towards LLM Empowered Recommendation via Tool Learning. SIGIR 2024.

Agent as Recommender

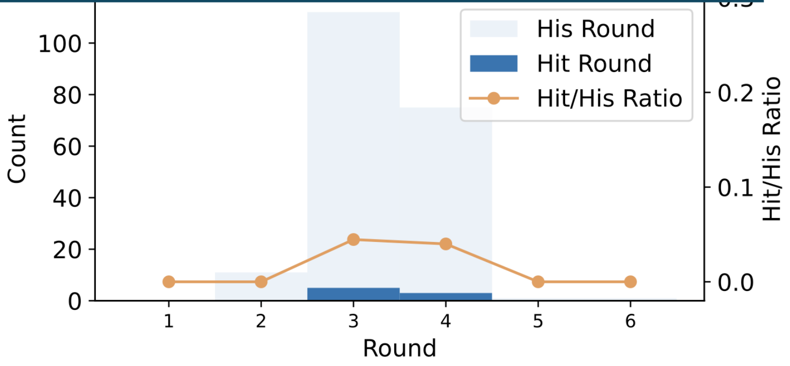
ToolRec

- **Key Observations:**
  - Benefiting from rank tools and tools, ToolRec excels on the ML-1M and Amazon-Book datasets compared to baseline recommenders, demonstrating that it can better align with the users' intent.

	ML-1M		Amazon-Book		Yelp2018	
	Recall	NDCG	Recall	NDCG	Recall	NDCG
SASRec	0.203±0.047	0.1017±0.016	0.047±0.015	0.0205±0.006	0.030±0.005	0.0165±0.006
ToolRec	**0.215**±0.044	**0.1171**±0.018	**0.053**±0.013	**0.0259**±0.005	0.028±0.003	0.0159±0.001
ToolRec$_B$	0.185±0.018	0.0895±0.002	0.043±0.013	0.0223±0.008	0.025±0.005	0.0136±0.009
Improvement	3.36%	15.10%	14.28%	5.14%	−29.16%	−27.32%



Agents Utilizing External Tools can Enhance Recommendations.



(b) Amazon-Book.

- ToolRec shows subpar performance on the Yelp2018 dataset - local (niche) businesses.
- Most processes conclude in three or four rounds, indicating that the LLM can understand user preferences after a few iterations.

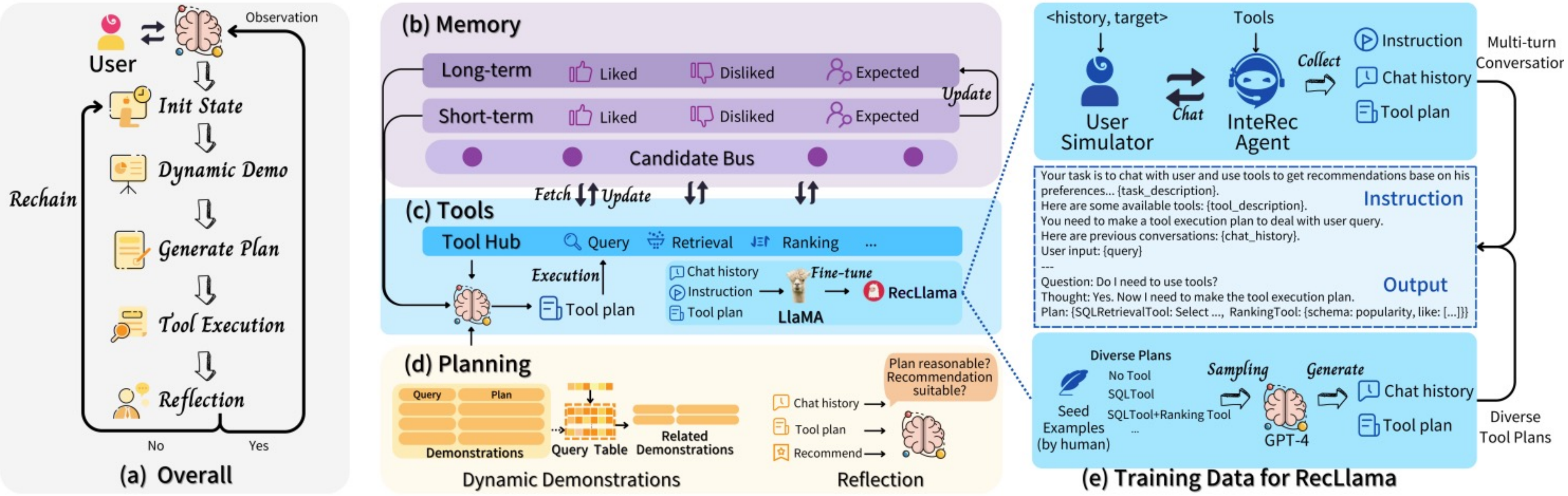
Yuyue Zhao et al. Let Me Do It For You: Towards LLM Empowered Recommendation via Tool Learning. SIGIR 2024.

Agent as Recommender
InteRecAgent

Agent as Recommender

❑ **InteRecAgent: Interactive Recommender.**

- **Key Points:**
  - Agents can create a versatile and interactive recommender system.



- InteRecAgent enables traditional recommender systems, such as those ID-based matrix factorization models, to become interactive systems with a natural language interface.

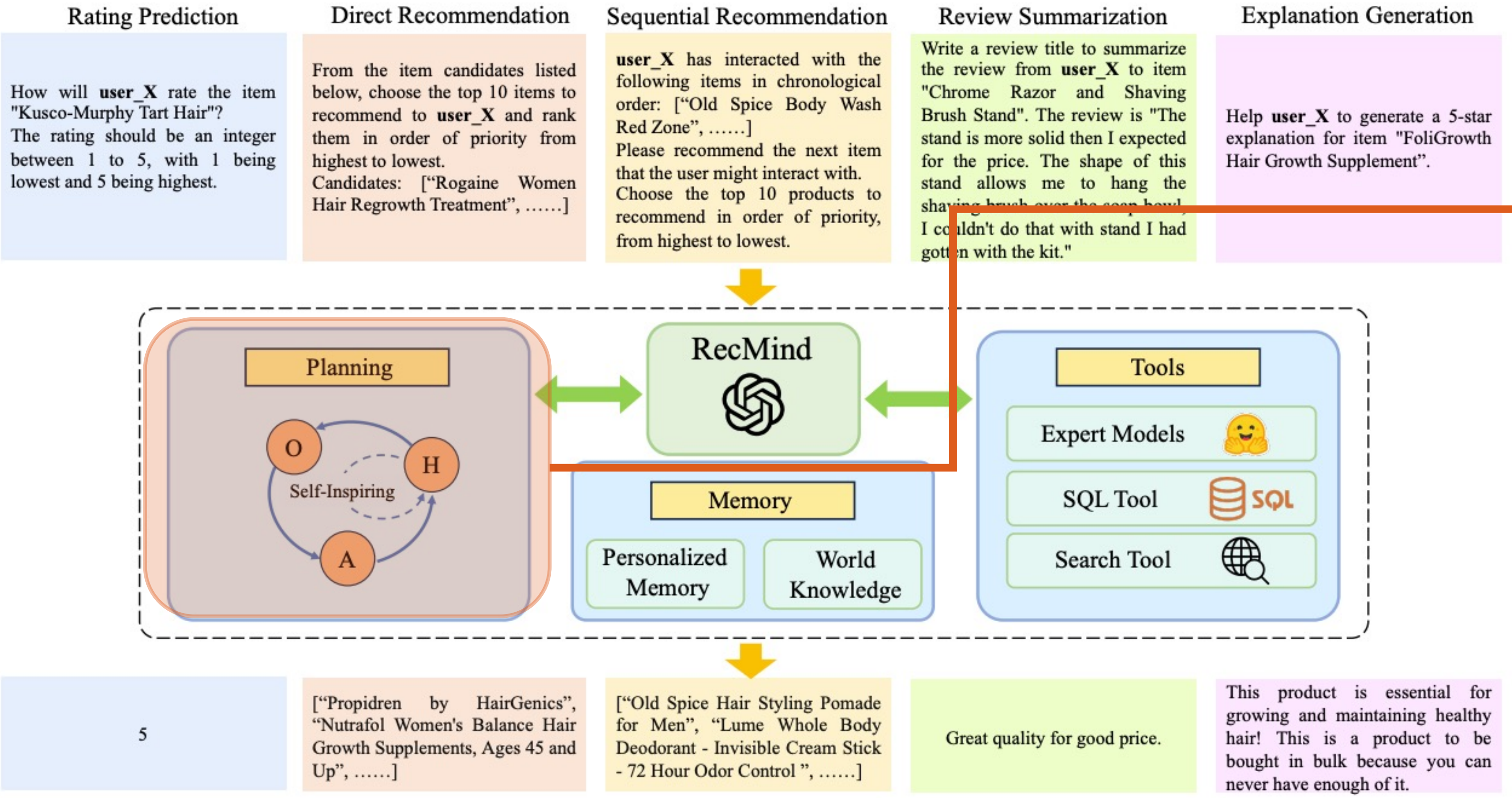
Xu Huang et al. Recommender AI Agent: Integrating Large Language Models for Interactive Recommendations. Arxiv 2023..

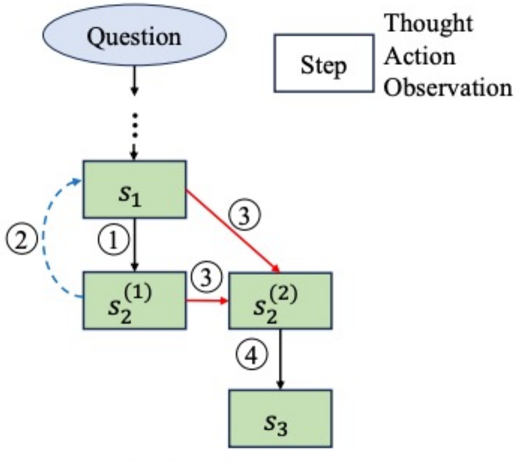
Agent as Recommender
RecMind

Agent as Recommender

- **Key Points:**
  - Can Agents with self-inspiring planning Enhance Recommendations?

❑ **RecMind: Recommender agent with Self-Inspiring planning ability**





(b) Self-Inspiring

- Self-inspires:
- At each intermediate planning step, the agent "self-inspires" to consider all previously explored paths for the next planning, both generating alternative thoughts and backtracking.

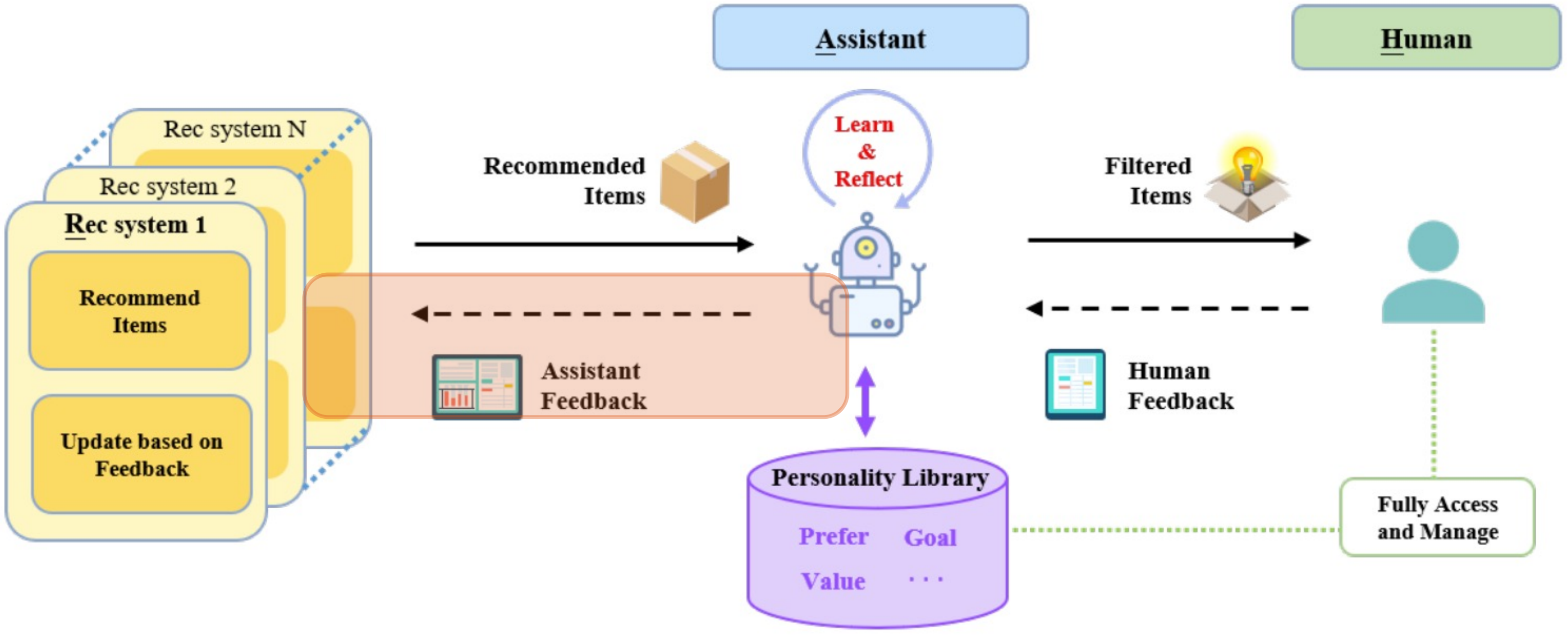
Yancheng Wang et al. "RecMind: Large Language Model Powered Agent For Recommendation.NAACL 2024.

Agent as Recommendation Assistant
RAH

Agent as Rec Assistant

❑ **RAH: Reflection-enhanced user alignment for Rec assistant**

- **Key Points:**
  - Can Agents with Learn-Act-Critic loop comprehend a user's personality from their behaviors?



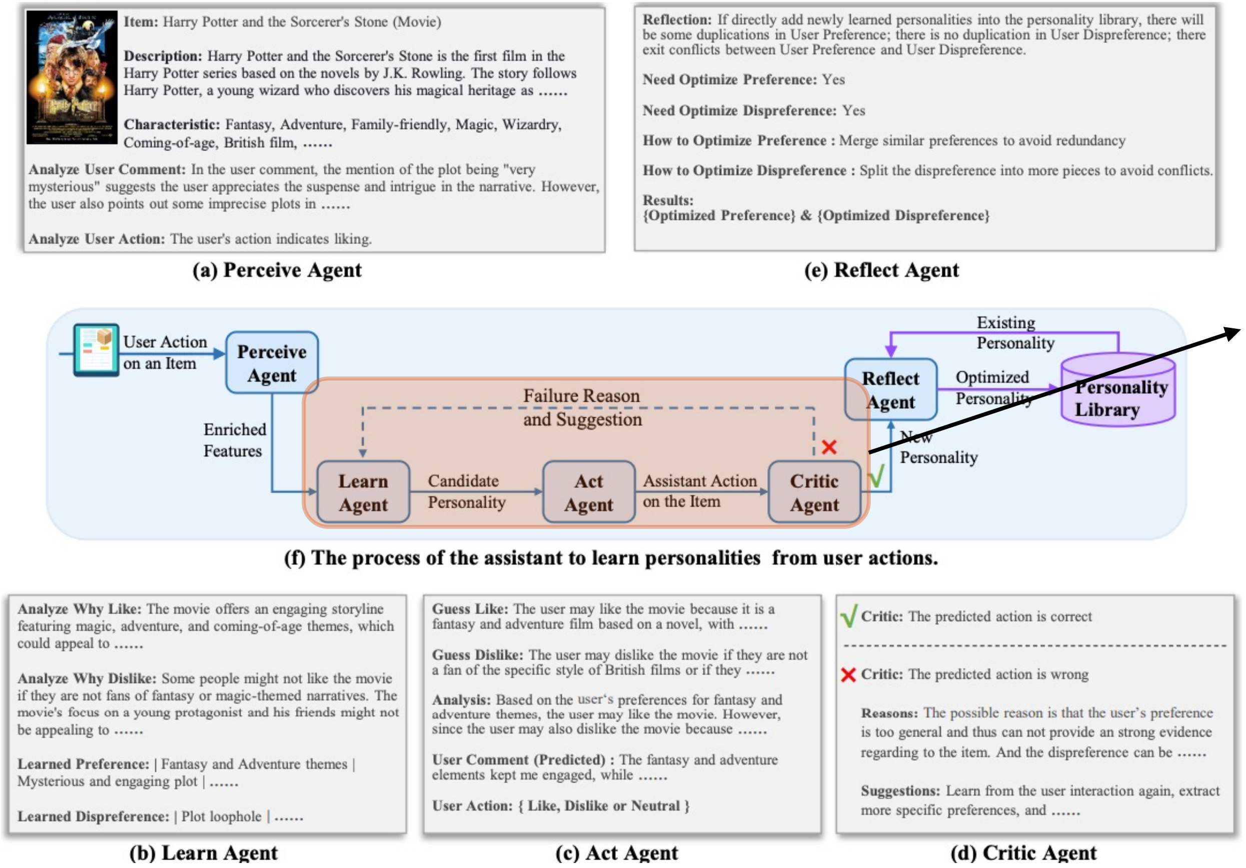
Yubo Shu et al. RAH! RecSys-Assistant-Human: A Human-Centered Recommendation Framework with LLM Agents. Arxiv 2023.

Agent as Recommendation Assistant
RAH

Agent as Rec Assistant

❑ **RAH: Reflection-enhanced user alignment for Rec assistant**

- **Key Points:**
  - Can Agents with Learn-Act-Critic loop comprehend a user's personality from their behaviors?



❖ Learn-Act-Critic Loop:
- Learn Agent collaborates with the Act and Critic Agents in an iterative process to grasp the user's personality.
- Upon receiving user feedback, Learn Agent extracts an <u>initial personality</u> as a candidate.
- Act Agent utilizes this candidate as input to <u>predict</u> the user's actual action.
- The Critic Agent then <u>assesses</u> the accuracy. If incorrect, Learn Agent <u>refines</u> the candidate's personality.

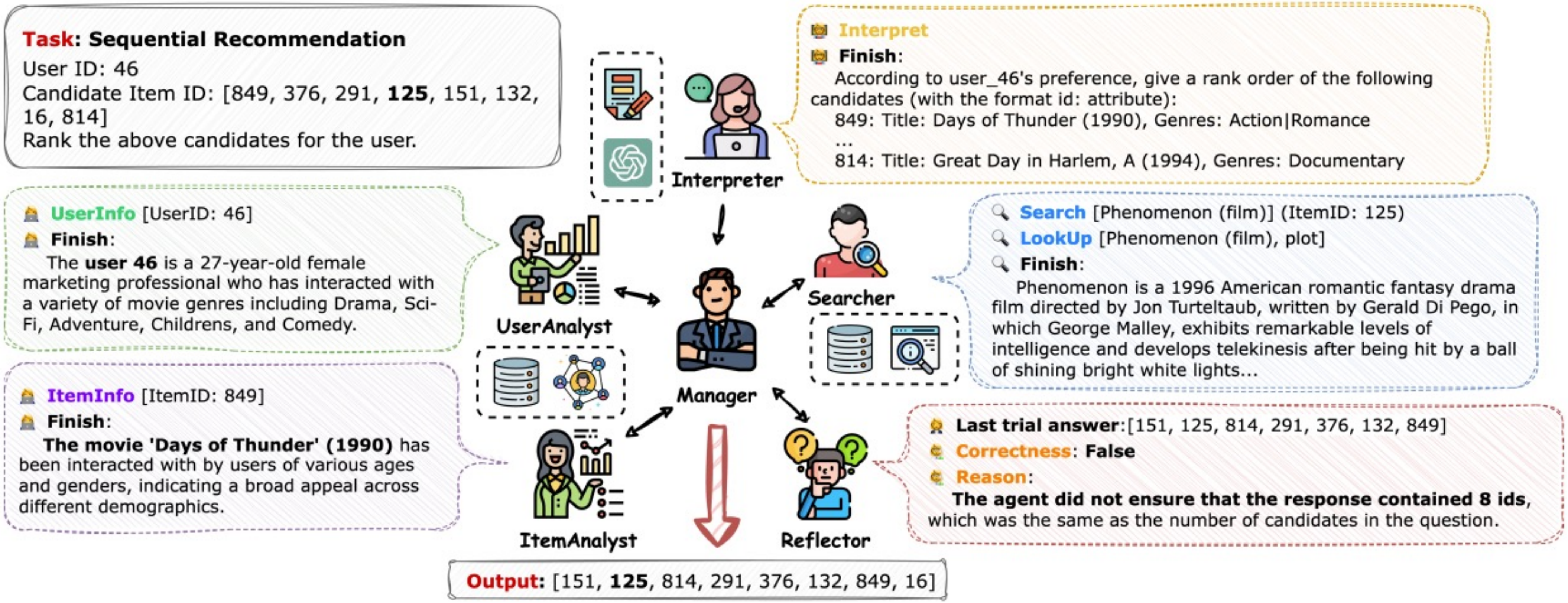
Yubo Shu et al. RAH! RecSys-Assistant-Human: A Human-Centered Recommendation Framework with LLM Agents. Arxiv 2023.

Multi-Agents as Recommender
MACRec

Multi-Agent as Recommender

❑ **MACRec: enhance RecSys through multi-agent collaboration**

- **Key Points:**
  - Multi-agents with different **roles** work collaboratively to tackle a specific recommendation task.



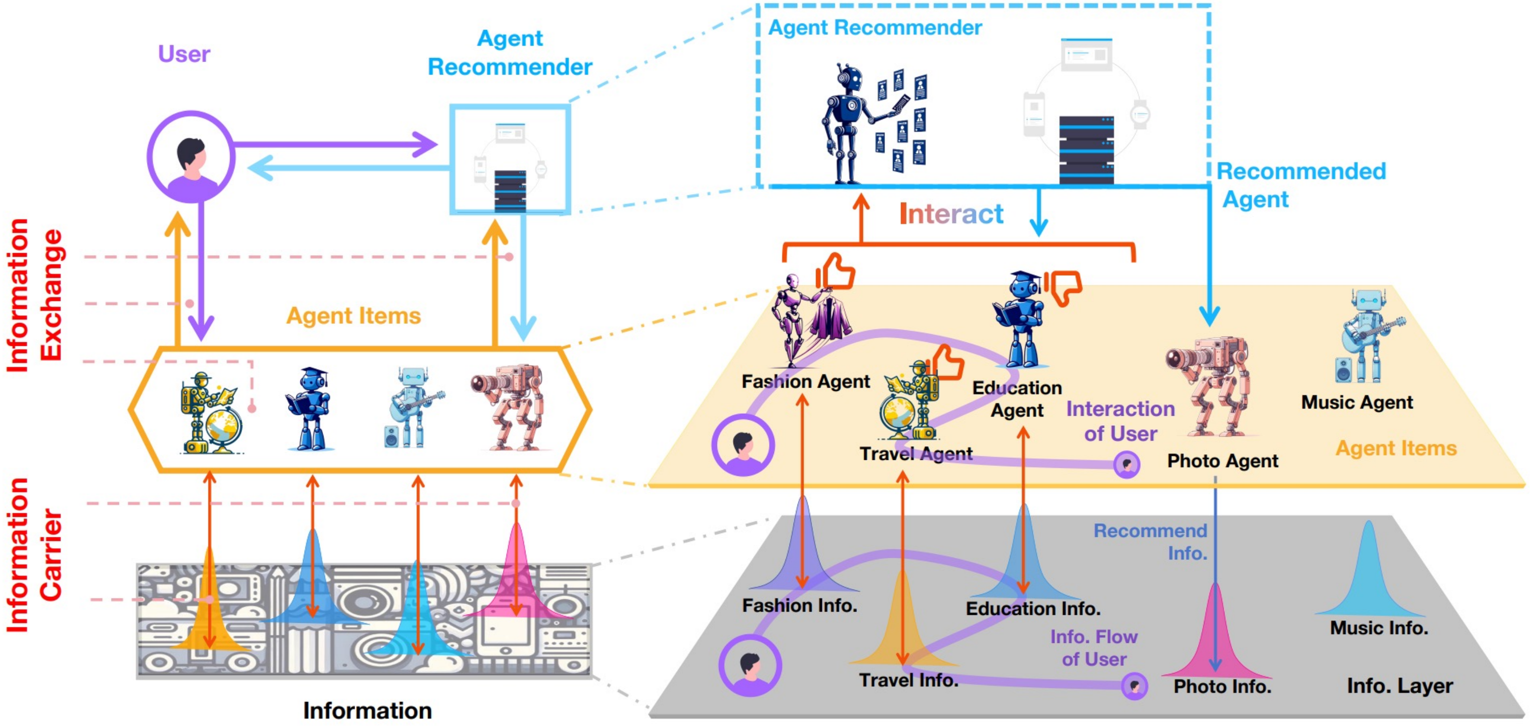
Zhefan Wang et al. Multi-Agent Collaboration Framework for Recommender Systems. Arxiv 2024.

Agent Recommender for Agent Platform
Rec4Agentverse

Agent Recommender

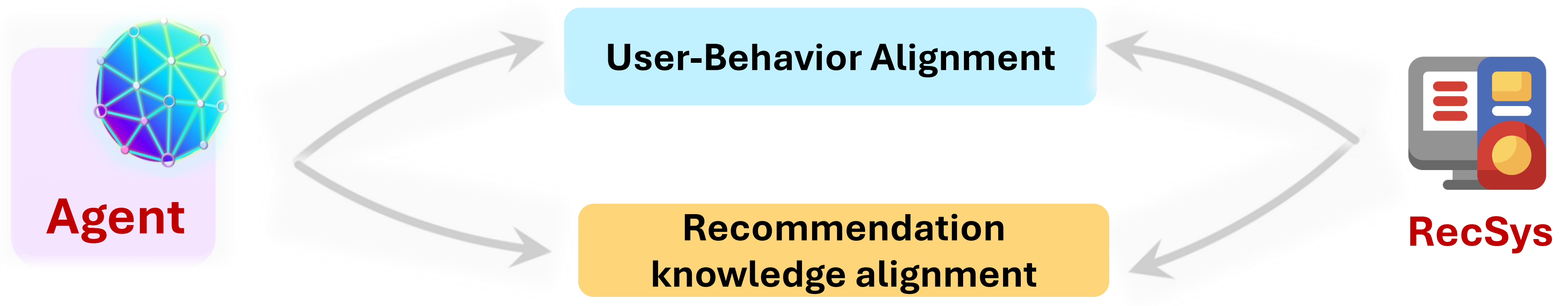
❑ **Rec4Agentverse: Agent recommender for Agent platform**

- **Key Points:**
  - Treating LLM-based Agents in Agent platform as items in the recommender system.
  - **Agent Recommender** is employed to recommend personalized Agent Items for each user.



Jizhi Zhange t al. Prospect Personalized Recommendation on Large Language Model-based Agent Platform. Arxiv 2024.

LLM-powered Agents in Recommendation



- **LLM-empowered have potentials to solve long-standing problems in recommendation**
  - Can an LLM-powered Agent faithfully simulate **users**?
    - **Agent4Rec, UGen, AgentCF, RecAgent**
  - Can an LLM-powered Agent be a better **recommender** with recommendation-specific knowledge?
    - **ToolRec, InteRecAgent, RecMind, RAH, MACRec, Rec4Agentverse**

Q&A

Thanks for listening!

Email: an_zhang@nus.edu.sg

An Zhang's Homepage

Resources