Toward High-Fidelity Multi-Agent Recommendation: An Agentic Design Framework Integrating RecoWorld and LLMs

Doctoral Consortium

Ionuț-Dragoș Neremzoiu
neremzoiu.ionut.k4c@student.ucv.ro
Department of Computers and Information Technology
University of Craiova, Faculty of Automatics, Computers and Electronics
Craiova, Romania

ABSTRACT

Traditional evaluation in Recommender Systems (RS) often relies on static, offline benchmarking, failing to capture the dynamic feedback loops between autonomous users and algorithmic policies. To address this, we propose an agentic design framework that treats recommendation as a Multi-Agent System (MAS). The framework integrates RecoWorld (environment), RecBole (system policy) and Large Language Models (reasoning agents), which enables a granular study of the co-evolutionary dynamics and emergent properties inherent in human-AI interaction. This doctoral work aims to provide a methodology for simulating long-term user retention, user-system co-evolution and assessing agent welfare.

KEYWORDS

LLM-based Recommendation Simulation; POMDP; RecBole; RecoWorld

ACM Reference Format:
Ionuț-Dragoș Neremzoiu. 2026. Toward High-Fidelity Multi-Agent Recommendation: An Agentic Design Framework Integrating RecoWorld and LLMs: Doctoral Consortium. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/HDBT9412

1 INTRODUCTION

Recent research has revealed a paradigm called *LLM as Recommendation Simulator* [8], which emphasizes LLMs as the foundational architecture of generative agents for the simulation of user behavior in the recommendation environment. Unlike traditional rule-based simulators, these autonomous agents, such as **InteRecAgent** [3], , **RecAgent** [7], **Agent4Rec**[9], and **AgentCF** [10], incorporate sophisticated **memory** and **reasoning modules**. This allows them to maintain latent user states (e.g., interests, satisfaction) and execute multi-step actions, which facilitates a high-fidelity simulation of user-system interaction trajectories.

In contrast to [3], [7], [9], and [10], which propose specific agentic recommenders, [5] introduces **RecoWorld** - a blueprint for creating simulated environments designed to accelerate the development of such systems. By emphasizing collective impacts, user simulation, and reward signal design, RecoWorld offers significant potential for enhancing generative models and improving the modeling of short- and long-term interests in existing recommender systems.

Given these premises, we introduce an agentic design framework that re-conceptualizes recommendation as a Multi-Agent System (MAS). By synthesizing the **RecoWorld** environment with its policy library and LLM-based reasoning agents, our framework facilitates the granular study of co-evolutionary dynamics and emergent properties within human-AI interactions. This doctoral research establishes a methodology for simulating long-term user retention, modeling user-system co-adaptation, and evaluating agent welfare over extended temporal horizons

2 FORMAL DEFINITION PROBLEM

We model the recommendation process as a **Partially Observable Markov Decision Process (POMDP)** [4], which allows us to account for the latent nature of user preferences and the sequential interaction between agents. The process is defined by the tuple $\langle S, A, T, R, \Omega, O \rangle$:

- S **(State Space):** The global state of the environment, managed by **RecoWorld**, which includes item metadata and the hidden, evolving preference profiles of the user population.
- A **(Action Space):** The set of possible recommendation lists generated by the **System Agent** using policies instantiated via **RecBole**.
- $T : S \times A \to \Pi(S)$ **(Transition Function):** The probability distribution over next states, representing how the environment evolves after a user interacts with a recommended item.
- $R : S \times A \to \mathbb{R}$ **(Reward Function):** The feedback signal (e.g., click, dwell time, or explicit rating) that guides the learning of the System Agent.
- Ω **(Observation Space):** The information perceived by the **User Agent**. In our framework, this is a semantic translation of the item features into natural language.
- $O : S \times A \to \Pi(\Omega)$ **(Observation Function):** The process of surfacing a subset of the global state S to the User Agent based on the System Agent's actions.
  We denote $\Pi(S)$ as the set of all probability distributions over the state space S.

The core research challenge is the **Alignment of Intelligence**: ensuring that the System Agent's algorithmic outputs (numerical embeddings) are semantically interpretable by the User Agent's cognitive reasoning module (the LLM).

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/HDBT9412

3 PROPOSED SYSTEM ARCHITECTURE

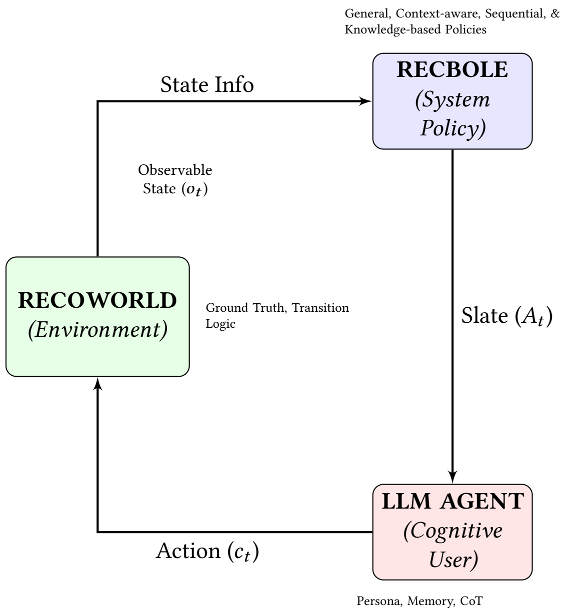

Figure 1: Agentic Design Framework: The closed-loop interaction between the Environment (RecoWorld), the System Agent (RecBole), and the Cognitive User Agent (LLM).

We propose a modular, three-tier architecture designed for extensibility and scale, which involves the following components: **the world model**, **the policy engine** and **the cognitive agent**.

3.1 The World Model

RecoWorld serves as the foundational environmental substrate for the simulation. It maintains the ground-truth document database and manages the global temporal state of the ecosystem. By operationalizing a Gym-like RL framework, the model facilitates standardized, multi-turn sequential interactions between the system and user agents. This architecture enables the generation of high-quality interaction trajectories, allowing for the extraction of engagement statistics that serve as "pseudo-reward" signals for training.

3.2 The Policy Engine (RecBole)

The System Agent is powered by RecBole. In our framework, RecBole is not an evaluator but a Policy Library. We leverage its optimized implementations of Neural Collaborative Filtering (NCF), Sequential Models (SASRec), and Graph-based models (LightGCN). This allows us to treat the Recommender as an autonomous entity with its own objectives (e.g., maximizing immediate engagement vs. long-term diversity).

3.3 The Cognitive User Agent (LLM)

The User Agent simulates realistic human behavior through an LLM-driven reasoning loop:

- **Persona and Context**: A profile integrating demographics, temporal factors, and history to define agent biases.
- **Reasoning and Perception**: A three-step *"Think it through"* process where the LLM evaluates slates against its mindset to deliberate across actions like clicking, liking, or skipping.
- **Instruction Management**: A feedback loop where the agent updates its internal state and generates reflective instructions (e.g., "show me more interesting content") to guide the system.

4 PRELIMINARY RESULTS: POLICY PERFORMANCE

To validate the Policy Engine, we benchmarked three **RecBole** models—**BPR** [6], **ConvNCF** [1], and **NeuMF** [2]—as the System Agent's initial strategies. We used the **ml-32m**[1] dataset (>32M interactions, 200,949 users, 84,433 items) with an 80/10/10 train/validation/test split. Evaluation utilized 50 negative samples per positive interaction across three top-K metrics (Table 1): **Recall@**K (retrieval coverage), **NDCG@**K (position-weighted ranking quality), and **MRR@**K (top-ranked accuracy). For this preliminary baseline, models were trained for a single epoch using one seed.

Table 1: Performance comparison of System Agent policies on the ml-32m dataset (Single-Epoch Training). All metrics are reported on the test set.

Model	Recall@10	Recall@20	NDCG@10	NDCG@20	MRR@10
ConvNCF	0.6794	0.8060	0.7962	0.8101	0.9024
BPR	0.6859	0.8112	0.8060	0.8198	0.9081
NeuMF	**0.7263**	**0.8479**	**0.8647**	**0.8752**	**0.9350**

The results indicate that **NeuMF** achieves the highest performance across all ranking-aware metrics even with minimal training iterations, providing a robust baseline for the System Agent's initial state. These scores serve as a prerequisite for the multi-turn interaction phase, where the **Cognitive User Agent** provides reflective instructions to further refine these policies in real-time.

5 CONCLUSION AND FUTURE WORK

We have introduced a novel agentic framework bridging static benchmarking with dynamic behavior modeling by integrating **RecoWorld** and **RecBole** [11]. Our approach advances evaluation through three key contributions: (1) **High-Fidelity Simulation**, where LLM-driven agents employ reasoning for multi-turn interactions; (2) **Instruction-Driven Feedback**, operationalizing natural language instructions within an RL framework; and (3) **Long-Term Optimization**, utilizing pseudo-rewards to prioritize retention over myopic engagement.

Future work will expand this into multi-agent social simulations to study collective impacts like interest drift. Ultimately, this framework establishes a foundation for collaborative systems where users and agents jointly shape sustainable information streams.

[1] https://grouplens.org/datasets/movielens/32m/

ACKNOWLEDGMENTS

I would like to express sincere gratitude to Prof. PhD Costin Bădică for his invaluable guidance, mentorship, and technical insights throughout the development of this research.

REFERENCES

[1] Xiangnan He, Xiaoyu Du, Xiang Wang, Feng Tian, Jinhui Tang, and Tat-Seng Chua. 2018. Outer Product-based Neural Collaborative Filtering. *ArXiv* abs/1808.03912 (2018). https://api.semanticscholar.org/CorpusID:51604671

[2] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. *Proceedings of the 26th International Conference on World Wide Web* (2017). https://api.semanticscholar.org/CorpusID:13907106

[3] Xu Huang, Jianxun Lian, Yuxuan Lei, Jing Yao, Defu Lian, and Xing Xie. 2023. Recommender AI Agent: Integrating Large Language Models for Interactive Recommendations. *ACM Transactions on Information Systems* 43, 2 (2023), 1–33.

[4] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. 1998. Planning and Acting in Partially Observable Stochastic Domains. *Artificial Intelligence* 101, 1-2 (1998), 99–134.

[5] Fei Liu, Xinyu Lin, Hanchao Yu, Mingyuan Wu, Jianyu Wang, Qiang Zhang, Zhuokai Zhao, Yinglong Xia, Yao Zhang, Weiwei Li, Mingze Gao, Qifan Wang, Lizhu Zhang, Benyu Zhang, and Xiangjun Fan. 2025. RecoWorld: Building Simulated Environments for Agentic Recommender Systems. *arXiv preprint arXiv:2509.10397* (2025).

[6] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. *ArXiv* abs/1205.2618 (2009). https://api.semanticscholar.org/CorpusID:10795036

[7] Lei Wang, Jingsen Zhang, Hao Yang, Zhi-Yuan Chen, Jiakai Tang, Zeyu Zhang, Xu Chen, Yankai Lin, Ruihua Song, Wayne Xin Zhao, Jun Xu, Zhicheng Dou, Jun Wang, and Ji-Rong Wen. 2025. User Behavior Simulation with Large Language Model-Based Agents. *ACM Transactions on Information Systems* 43, 2 (2025), 1–37.

[8] Lanling Xu, Junjie Zhang, Bingqian Li, Jinpeng Wang, Sheng Chen, Wayne Xin Zhao, and Ji-Rong Wen. 2024. Tapping the Potential of Large Language Models as Recommender Systems: A Comprehensive Framework and Empirical Analysis. *ACM Transactions on Knowledge Discovery from Data* 19 (2024), 1 – 51. https://api.semanticscholar.org/CorpusID:266902921

[9] An Zhang, Leheng Sheng, Yuxin Chen, Hao Li, Yang Deng, Xiang Wang, and Tat-Seng Chua. 2024. On Generative Agents in Recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1807–1817.

[10] Junjie Zhang, Yupeng Hou, Ruobing Xie, Wenqi Sun, Julian McAuley, Wayne Xin Zhao, Leyu Lin, and Ji-Rong Wen. 2024. AgentCF: Collaborative Learning with Autonomous Language Agents for Recommender Systems. In *Proceedings of the ACM Web Conference 2024*. 3679–3689.

[11] Wayne Xin Zhao, Shanlei Mu, Yupeng Hou, Zihan Lin, Kaiyuan Li, Yushuo Chen, Yujie Lu, Haibo Wang, Changxin Tian, Xingyu Pan, Yingqian Min, Zhichao Feng, Xinyan Fan, Xu Chen, Pengfei Wang, Wendi Ji, Yaliang Li, Xiaoling Wang, and Ji rong Wen. 2020. RecBole: Towards a Unified, Comprehensive and Efficient Framework for Recommendation Algorithms. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (2020). https://api.semanticscholar.org/CorpusID:226237553