A Literature Review on Simulation in Conversational Recommender Systems

Haoran Zhang
Tianjin University
zhr00@tju.edu.cn

Xin Zhao
University of Electronic Science and Technology of China
zxin11@std.uestc.edu.cn

Jinze Chen
University of Washington
jzch777@uw.edu

Junpeng Guo
Tianjin University
guojp@tju.edu.cn

Abstract

The emergence of pre-trained language models has rapidly attracted significant research interest in Conversational Recommender Systems (CRSs) and accelerated their development. While CRSs' complexity in Human-Computer Interaction (HCI) offers new opportunities for accurate user modeling and item recommendations, it also presents challenges in data and system evaluation. Simulation rises as a promising approach to address these challenges. This study develops a taxonomy framework to systematically review CRSs and their simulation studies, identifying key research challenges and opportunities. Our review not only consolidates existing knowledge but also uncovers understudied areas, providing a roadmap for future research in CRSs and HCI.

Keywords: Conversational recommender systems, Simulation, Human-computer interaction, Literature review

This document is a complete list of references to the submission.

References

1. Afzali, J., Drzewiecki, A. M., Balog, K., & Zhang, S. (2023). UserSimCRS: A User Simulation Toolkit for Evaluating Conversational Recommender Systems. *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, 1160-1163. https://doi.org/10.1145/3539597.3573029
2. Balog, K., & Zhai, C. (2024). User Simulation for Evaluating Information Access Systems. *Foundations and Trends® in Information Retrieval*, 18(1-2), Article 1-2. https://doi.org/10.1561/1500000098
3. Bernard, N., & Balog, K. (2024). Identifying Breakdowns in Conversational Recommender Systems using User Simulation. *ACM Conversational User Interfaces 2024*, 1-10. https://doi.org/10.1145/3640794.3665539
4. Di Bratto, M., Origlia, A., Di Maro, M., & Mennella, S. (2024). Linguistics-based dialogue simulations to evaluate argumentative conversational recommender systems. *User Modeling and User-Adapted Interaction*, 34(5), Article 5. https://doi.org/10.1007/s11257-024-09403-3
5. Gao, C., Lei, W., He, X., De Rijke, M., & Chua, T.-S. (2021). Advances and challenges in conversational recommender systems: A survey. *AI Open*, 2, 100-126. https://doi.org/10.1016/j.aiopen.2021.06.002

6. Jannach, D. (2023). Evaluating conversational recommender systems: A landscape of research. *Artificial Intelligence Review*, 56(3), Article 3. https://doi.org/10.1007/s10462-022-10229-x
7. Jannach, D., & Chen, L. (2022). Conversational recommendation: A grand AI challenge. *AI Magazine*, 43(2), Article 2. https://doi.org/10.1002/aaai.12059
8. Jannach, D., Manzoor, A., Cai, W., & Chen, L. (2022). A Survey on Conversational Recommender Systems. *ACM Computing Surveys*, 54(5), Article 5. https://doi.org/10.1145/3453154
9. Jin, Y., Chen, L., Cai, W., & Zhao, X. (2024). CRS-Que: A User-centric Evaluation Framework for Conversational Recommender Systems. *ACM Transactions on Recommender Systems*, 2(1), Article 1. https://doi.org/10.1145/3631534
10. Li, C., Hu, H., Zhang, Y., Kan, M.-Y., & Haizhou Li. (2023). A Conversation is Worth A Thousand Recommendations: A Survey of Holistic Conversational Recommender Systems. *Proceedings of the Fifth Knowledge-Aware and Conversational Recommender Systems Workshop Co-Located with 17th ACM Conference on Recommender Systems (RecSys 2023)*, 3560, 7-20.
11. Li, R., Ebrahimi Kahou, S., Schulz, H., Michalski, V., Charlin, L., & Pal, C. (2018). Towards Deep Conversational Recommendations. *Advances in Neural Information Processing Systems*, 31.
12. Liang, T., Jin, C., Wang, L., Fan, W., Xia, C., Chen, K., & Yin, Y. (2024). LLM-REDIAL: A Large-Scale Dataset for Conversational Recommender Systems Created from User Behaviors with LLMs. *Findings of the Association for Computational Linguistics ACL 2024*, 8926-8939. https://doi.org/10.18653/v1/2024.findings-acl.529
13. Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative Agents: Interactive Simulacra of Human Behavior. Proceedings of the 36th Annual *ACM Symposium on User Interface Software and Technology*, 1-22. https://doi.org/10.1145/3586183.3606763
14. Pramod, D., & Bafna, P. (2022). Conversational recommender systems techniques, tools, acceptance, and adoption: A state of the art review. *Expert Systems with Applications*, 203, 117539. https://doi.org/10.1016/j.eswa.2022.117539
15. Shanahan, M., McDonell, K., & Reynolds, L. (2023). Role play with large language models. *Nature*, 623(7987), 493-498. https://doi.org/10.1038/s41586-023-06647-8
16. Shen, Q., Wu, L., Zhang, Y., Pang, Y., Wei, Z., Xu, F., Long, B., & Pei, J. (2024). Multi-Interest Multi-Round Conversational Recommendation System with Fuzzy Feedback Based User Simulator. *ACM Transactions on Recommender Systems*, 2(4), Article 4. https://doi.org/10.1145/3616379
17. Suglia, A., Greco, C., Basile, P., Semeraro, G., & Caputo, A. (2017). An Automatic Procedure for Generating Datasets for Conversational Recommender Systems. *Working Notes of CLEF 2017*, 1866.
18. Vlachou, M., & Macdonald, C. (2024). What Else Would I Like? A User Simulator using Alternatives for Improved Evaluation of Fashion Conversational Recommendation Systems. *arXiv*. https://doi.org/10.48550/arXiv.2401.05783
19. Wang, L., Zhang, J., Yang, H., Chen, Z.-Y., Tang, J., Zhang, Z., Chen, X., Lin, Y., Sun, H., Song, R., Zhao, X., Xu, J., Dou, Z., Wang, J., & Wen, J.-R. (2025). User Behavior Simulation with Large Language Model-based Agents. *ACM Transactions on Information Systems*, 43(2), 1-37. https://doi.org/10.1145/3708985

20. Wang, X., Tang, X., Zhao, X., Wang, J., & Wen, J.-R. (2023). Rethinking the Evaluation for Conversational Recommendation in the Era of Large Language Models. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 10052-10065. https://doi.org/10.18653/v1/2023.emnlp-main.621
21. Wang, X., Zhou, K., Tang, X., Zhao, W. X., Pan, F., Cao, Z., & Wen, J.-R. (2023). Improving Conversational Recommendation Systems via Counterfactual Data Simulation. *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2398-2408. https://doi.org/10.1145/3580305.3599387
22. Yoon, S., He, Z., Echterhoff, J., & McAuley, J. (2024). Evaluating Large Language Models as Generative User Simulators for Conversational Recommendation. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, 1490-1504. https://doi.org/10.18653/v1/2024.naacl-long.83
23. Zhang, E., Wang, X., Gong, P., Lin, Y., & Mao, J. (2024). USimAgent: Large Language Models for Simulating Search Users. *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2687-2692. https://doi.org/10.1145/3626772.3657963
24. Zhang, J., Adomavicius, G., Gupta, A., & Ketter, W. (2020). Consumption and Performance: Understanding Longitudinal Dynamics of Recommender Systems via an Agent-Based Simulation Framework. *Information Systems Research*, 31(1), 76-101. https://doi.org/10.1287/isre.2019.0876
25. Zhang, S., & Balog, K. (2020). Evaluating Conversational Recommender Systems via User Simulation. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1512-1520. https://doi.org/10.1145/3394486.3403202
26. Zhang, S., Wang, M.-C., & Balog, K. (2022). Analyzing and Simulating User Utterance Reformulation in Conversational Recommender Systems. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 133-143. https://doi.org/10.1145/3477495.3531936
27. Zhang, Y., Chen, X., Ai, Q., Yang, L., & Croft, W. B. (2018). Towards Conversational Search and Recommendation: System Ask, User Respond. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 177-186. https://doi.org/10.1145/3269206.3271776
28. Zhu, L., Huang, X., & Sang, J. (2024). How Reliable is Your Simulator? Analysis on the Limitations of Current LLM-based User Simulators for Conversational Recommendation. *Companion Proceedings of the ACM Web Conference 2024*, 1726-1732. https://doi.org/10.1145/3589335.3651955