

# Report on The 1st Workshop on Human-Centered Recommender Systems (HCRS 2025) at TheWebConf 2025

Kaike Zhang

University of Chinese Academy of Sciences  
China  
kaikezhang99@gmail.com

Yunfan Wu

University of Chinese Academy of Sciences  
China  
wuyunfan666@gmail.com

Youngang Lyu, Du Su, Yingqiang Ge, Shuchang Liu, Qi Cao, Zhaochun Ren,  
Fei Sun\*

## Abstract

This report summarizes the 1st Workshop on Human-Centered Recommender Systems (HCRS), held on April 29, 2025, in conjunction with TheWebConf 2025. Recommender systems have become quintessential applications of human-computer interaction, offering significant convenience in daily life while simultaneously presenting challenges such as the information cocoon effect, privacy concerns, and fairness issues. Consequently, this workshop was established to provide a platform for researchers to explore the development of recommender systems that prioritize human needs, values, and capabilities at the core of their design and operation. Topics covered included robustness, privacy, transparency, fairness, diversity, accountability, ethical considerations, and user-friendly design. Through these discussions, the workshop sought to foster a collaborative environment for sharing insights on implementing these properties and exploring innovative evaluation metrics, ultimately advancing the field toward more ethical, user-centric, and socially responsible recommender systems.

**Date:** 29 April 2025.

**Website:** <https://human-centeredrec.github.io/>.

## 1 Introduction

Recommender systems serve as essential tools for managing the exponential growth of information online, influencing users' decisions across domains such as e-commerce, media consumption, and social networking. However, their pervasive use has introduced significant challenges, including the information cocoon effect, privacy breaches, and fairness concerns. To address these issues, there is a growing consensus on the necessity of developing Human-Centered Recommender Systems (HCRS) [Konstan and Terveen, 2021; Silva et al., 2024]. HCRS refers to the development of

---

\*Affiliation not shown for all authors due to space limitations (see Appendix A for details).

---

systems that prioritize human needs, values, and capabilities at the core of their design and operation.

The workshop was convened on April 29, 2025, at The Web Conference (TheWebConf) 2025. Since the invention of the World Wide Web in 1989, TheWebConf (formerly WWW) has been the premier international venue for discussing the future direction of the Web. By assembling scholars, practitioners, and end-users with the goal of envisioning the future of the Web, TheWebConf provided the ideal ecosystem for discussing the human-centric evolution of recommender technologies.

A critical discussion point within the workshop was distinguishing HCRS from related paradigms. While concepts like Trustworthy Recommender Systems and Responsible Recommender Systems share similarities with HCRS in addressing ethical and societal concerns, they differ in their primary focus. Trustworthy Recommender Systems aim to build systems that users can rely on, emphasizing technical aspects such as transparency, robustness, privacy, and security [Ge et al., 2022; Wang et al., 2024; Zhang et al., 2025a]. Responsible Recommender Systems typically concentrate on aligning system operations with ethical principles and social responsibilities, often focusing on fairness, accountability, and mitigating biases [Kazienko and Cambria, 2024]. In contrast, HCRS places the human user at the center of the design process. It strives to enhance user well-being, satisfaction, and empowerment by deeply considering human values and experiences, actively involving users in the design and evaluation loop to ensure systems align with their goals and capabilities [Konstan and Terveen, 2021; Shneiderman, 2020; Lowdermilk, 2013].

## 2 Scope and Topics

The workshop provided a collaborative platform for researchers to present the latest advancements in the rapidly evolving field of HCRS. By fostering interdisciplinary dialogue and knowledge exchange, the event aimed to advance the understanding and implementation of recommender systems that are not only technically effective but also ethically sound, user-centered, and socially responsible. The workshop encompassed original research contributions focusing on a diverse range of topics, including:

- **Robustness:** Fraudster Detection, Defense Against Adversarial Attacks, Vulnerabilities of LLM-based RS, Certifiable Robustness, Denoising, Data Sparsity, Cross-Domain Robustness, User-Aware Robustness, etc.
- **Privacy:** Differential Privacy, Federated Learning in RS, Data Ownership, Privacy Risks in LLM-based RS, Data Anonymization, Membership Inference Attack, Data Minimization, Unlearning, etc.
- **Transparency:** Explainable RS, Interpretable RS, User-Centric Explanation Generation, Causal Explanations, Neurosymbolic Reasoning for RS, LLMs for Transparent RS, etc.
- **Fairness and Bias:** Debiasing in RS, Fairness and Bias in LLM-based RS, etc.
- **Diversity:** Content Diversity, Recommendation Diversity, User Perceptions of Diversity and Personalization, Addressing Filter Bubbles and Echo Chambers, etc.
- **Ethics of Recommender Systems:** Ethical Frameworks, Mitigation of Misinformation Spread, Ethical Implications of Personalized Content, User Consent and Ethical Data Usage, Ethical Challenges in LLM-based RS, etc.

- 
- **Accountability:** Traceability, Responsible RS, Controllable Recommendations, etc.
  - **Human-Computer Interaction Design:** User Interface Design, Interactive and Conversational RS, Accessibility and Inclusivity, etc.
  - **Evaluation, Auditing, and Governance:** Evaluation Metrics, User Studies, Algorithm Auditing, Simulation, Governance Models, etc.

## 3 Rationale

The rapid advancement of human-centered recommender systems has catalyzed a wave of innovative research efforts in recent months [Wu et al., 2024; Wang et al., 2024; Deldjoo et al., 2024; Shen et al., 2025; Balloccu et al., 2024]. Consequently, this workshop was organized to address the critical need for a dedicated platform to foster discussions on these evolving directions.

### 3.1 Objectives and Outcomes

The primary objective of the workshop was to encourage researchers to propose new theoretical frameworks, interdisciplinary approaches, and perspectives for human-centered recommender systems. Discussions advocated for the adoption of advanced technologies, such as large language models, to enhance the human-centric qualities of existing systems. Furthermore, participants explored the development of innovative evaluation frameworks and metrics tailored to assessing these qualities. Looking forward, the insights generated from this workshop are expected to drive the evolution of recommender systems into broader domains, fostering their applicability across various contexts and contributing positively to societal well-being.

### 3.2 Diversity and Inclusion

**Diversity among Organizers.** The organizing committee was characterized by significant diversity in terms of gender, affiliations, countries of origin, scientific backgrounds, and levels of seniority. This diverse composition enhanced the capacity to consider a wide range of perspectives, ultimately enriching the discussions and outcomes of the workshop.

**Diversity among Speakers.** The workshop featured keynote talks from researchers with varied academic backgrounds, ensuring a broad spectrum of insights and expertise. The program showcased submissions from scholars worldwide, highlighting diverse voices within the research community.

To promote diversity within the panel discussions, the workshop curated a mix of participants, including both academics and industry professionals, representing various levels of seniority and experience. This approach fostered rich dialogue, allowing for a comprehensive exchange of ideas and perspectives that reflected the multifaceted nature of human-centered recommender systems.

## 4 Workshop Report

The 1st Workshop on Human-Centered Recommender Systems (HCRS) [Zhang et al., 2025c] served as a premier venue for uniting researchers and practitioners to discuss the human aspects

---

of recommendation technologies. The event attracted substantial interest, with approximately 40 in-person participants and over 60 remote attendees via Zoom. The inaugural edition received 14 submissions, of which 6 were accepted, reflecting a high standard of quality and relevance to the emerging field of human-centric AI.

## 4.1 Keynotes

### On Human Preference Alignment for LLM-based Recommenders

*Speaker: Professor Xiang Wang*

Professor Xiang Wang addressed the emerging frontier of human-centric preference alignment in large language models (LLMs). The talk focused on ensuring that model outputs align with human values and preferences within recommendation scenarios. Professor Wang introduced the S-DPO framework, a novel approach that effectively integrates preference alignment with recommendation-specific ranking objectives. The presentation demonstrated how this method instills ranking information into the language model, enabling LLM-based recommenders to distinguish preferred items from negative samples rather than focusing solely on positives. Furthermore, the talk outlined an extension of the framework that incorporates a self-improving mechanism for continuous adaptation.

### From Recommendations to Interactions: Putting Users Back in the Loop

*Speaker: Professor Aixin Sun*

Professor Aixin Sun provided a critical perspective on the user’s role in the interaction process. He highlighted that while considerable effort has been devoted to learning preferences from historical data, the user’s active role—particularly the phase between recommendation presentation and observed interaction—remains underexplored. Professor Sun delved into the complexities of user-item interactions, including decision-making costs, multi-step engagements, and unobservable behaviors that influence model design. The keynote concluded with a discussion on the trade-offs between task specificity and model generalizability, emphasizing that well-defined task formulations are foundational for robust evaluation.

## 4.2 Paper Presentations

**Filtering Discomforting Recommendations with Large Language Models** [Liu et al., 2025]: This work addresses the challenge of personalized algorithms inadvertently exposing users to subjectively discomforting content. The authors introduced *DiscomfortFilter*, an LLM-based tool that constructs a comprehensible and editable preference profile. By allowing users to mask discomforting preferences through natural language conversation, the tool filters recommendations in a plug-and-play manner without requiring access to black-box recommendation algorithms, thereby enhancing user control and mitigating negative emotional consequences.

**InstructAgent: Building User Controllable Recommender via LLM Agent** [Xu et al., 2025]: Critiquing the traditional user-platform paradigm where users lack control, this paper proposes a novel User-Agent-Platform paradigm. The authors presented *InstructAgent* and *Instruct<sup>2</sup>Agent*, which utilize Large Language Models as a protective shield between the user and the system. These agents interpret free-text user instructions and employ a dynamic memory

---

mechanism to optimize recommendations based on individual feedback, effectively mitigating echo chambers and reducing model bias against less-active users.

**Revealing, Analyzing, and Mitigating the Adverse Impact of Shared Accounts on Personalized Recommendations for Single-User Accounts** [Jing et al., 2025]: This study reveals that shared accounts (e.g., families sharing a TV) create spurious item-item associations that degrade recommendation performance for single-user accounts by 10%-40%. Through metric analysis, the authors attributed this to the distortion of item embedding proximity. To address this, they proposed a mitigation strategy leveraging multimodal item information to construct a reliable semantic graph, which counterbalances the noisy collaborative signals introduced by shared account behaviors.

**Improving LLM-Based Recommender Systems with User-Controllable Profiles** [Woźniak et al., 2025]: Focusing on user agency, this research investigates the efficacy of LLMs in capturing user preferences through editable natural language profiles. The authors compared different input representations—user profiles, few-shot history, and combined approaches—across various open-source and proprietary LLMs. Their findings demonstrate that incorporating user-controllable profiles significantly enhances recommendation performance compared to relying solely on historical data, offering a pathway toward more transparent and adaptable systems.

**On the Evaluation of Session-based Recommendation Unlearning** [Xin et al., 2024]: This paper tackles the challenge of "unlearning" specific data in session-based recommendation systems given sequential dependencies. The authors proposed the *SRU* framework, utilizing similarity-based partitioning and attentive aggregation to enable efficient unlearning. Additionally, the presentation introduced three data deletion strategies and novel evaluation metrics—including membership inference attacks and hit ratios—to rigorously verify the effectiveness of the unlearning process at both item and session levels.

**Evaluating Conversational Recommender Systems with Large Language Models: A User-Centric Evaluation Framework** [Chen et al., 2025]: To address the complexity of evaluating Conversational Recommender Systems (CRS), this work proposes a comprehensive user-centric evaluation framework powered by LLMs. The framework employs LLM-based user simulators to generate interaction logs and evaluates systems across four dimensions: Dialogue Behavior, Language Expression, Recommendation Items, and Response Content. Utilizing 12 distinct metrics, the study provides a standardized method to assess user satisfaction, effectively bridging the gap between human-computer interaction and recommender system evaluation.

### 4.3 2nd HCRS Announcement

Following the success of the inaugural event, we are pleased to announce that the 2nd HCRS workshop will be held at The Web Conference (WWW) 2026 [Zhang et al., 2025b]. We warmly invite all researchers and practitioners to participate and submit their latest work to continue this important dialogue. For more information, please visit our website at link<sup>1</sup>.

---

<sup>1</sup><https://hcrec.github.io/>

---

## 5 Conclusion and Outlook

The successful conclusion of the 1st Workshop on Human-Centered Recommender Systems (HCRS) marks a significant milestone in the evolution of recommendation technologies. By convening at TheWebConf 2025, the workshop successfully established a dedicated forum for shifting the research focus from system-centric optimization to human-centric value realization.

The diversity of the presented works—ranging from LLM-based preference alignment and user agents to unlearning frameworks and shared account mitigation—underscores the multifaceted nature of this emerging field. A recurring theme observed throughout the keynotes and paper presentations was the pivotal role of Large Language Models (LLMs) in bridging the gap between algorithmic complexity and user intent. The discussions highlighted that achieving true human-centeredness requires not only robust and fair algorithms but also mechanisms that grant users greater agency, transparency, and control over their digital experiences.

Despite the progress demonstrated, several challenges remain. The community identified a critical need for standardized evaluation protocols that can quantitatively measure abstract concepts such as user well-being, trust, and long-term satisfaction. Furthermore, as recommender systems become more conversational and interactive, the boundaries between recommendation and general information retrieval continue to blur, necessitating interdisciplinary collaboration between AI, HCI, and social science researchers.

We are confident that the momentum generated by this inaugural workshop will drive meaningful advancements in the coming years. We look forward to seeing how the ideas sparked here will mature and evolve, and we eagerly anticipate reconvening with the community at the 2nd HCRS workshop in 2026.

## Acknowledgments

We would like to express our sincere gratitude to the organizers of TheWebConf 2025 for their unwavering support and for providing an excellent platform to host this inaugural workshop. We are deeply indebted to our distinguished keynote speakers, Professor Xiang Wang and Professor Aixun Sun, whose insightful presentations set a high standard for our discussions and inspired the community.

Our heartfelt thanks go to the members of the Program Committee for their time and dedication in reviewing the submissions and providing constructive feedback, which was instrumental in ensuring the high quality of the workshop program. We also thank all the authors for sharing their innovative work and the participants for their active engagement and stimulating discussions. Finally, we acknowledge the collective effort of the Human-Centered Recommender Systems community in driving this important research agenda forward.

## A Authors and Affiliations

This is a list of authors of this publication. Workshop organizers:

- Kaike Zhang, University of Chinese Academy of Sciences, China
- Yunfan Wu, University of Chinese Academy of Sciences, China

- 
- Yougang Lyu, University of Amsterdam, Netherlands
  - Du Su, State Key Laboratory of AI Safety, ICT, CAS, China
  - Yingqiang Ge, Amazon, USA
  - Shuchang Liu, Kuaishou, China
  - Qi Cao, State Key Laboratory of AI Safety, ICT, CAS, China
  - Zhaochun Ren, Leiden University, ICT, CAS, Netherlands
  - Fei Sun, State Key Laboratory of AI Safety, ICT, CAS, China

## References

- Giacomo Balloccu, Ludovico Boratto, Gianni Fenu, Francesca Maridina Mallocci, and Mirko Marras. Explainable recommender systems with knowledge graphs and language models. In *European Conference on Information Retrieval*, pages 352–357. Springer, 2024.
- Nuo Chen, Quanyu Dai, Xiaoyu Dong, Xiao-Ming Wu, and Zhenhua Dong. Evaluating conversational recommender systems with large language models: A user-centric evaluation framework. *arXiv e-prints*, pages arXiv–2501, 2025.
- Yashar Deldjoo, Dietmar Jannach, Alejandro Bellogin, Alessandro Difonzo, and Dario Zanzonelli. Fairness in recommender systems: Research landscape and future directions. *User Modeling and User-Adapted Interaction*, 34(1):59–108, 2024.
- Yingqiang Ge, Shuchang Liu, Zuohui Fu, Juntao Tan, Zelong Li, Shuyuan Xu, Yunqi Li, Yikun Xian, and Yongfeng Zhang. A survey on trustworthy recommender systems. *ACM Transactions on Recommender Systems*, 2022.
- Xingbo Jing, Tao Lian, and Li Wang. Revealing, analyzing, and mitigating the adverse impact of shared accounts on personalized recommendations for single-user accounts. In *Companion Proceedings of the ACM on Web Conference 2025*, pages 2092–2101, 2025.
- Przemysław Kazienko and Erik Cambria. Toward responsible recommender systems. *IEEE Intelligent Systems*, 39(3):5–12, 2024.
- Joseph Konstan and Loren Terveen. Human-centered recommender systems: Origins, advances, challenges, and opportunities. *AI Magazine*, 42(3):31–42, 2021.
- Jiahao Liu, Yiyang Shao, Peng Zhang, Dongsheng Li, Hansu Gu, Chao Chen, Longzhi Du, Tun Lu, and Ning Gu. Filtering discomforting recommendations with large language models. In *Proceedings of the ACM on Web Conference 2025*, pages 3639–3650, 2025.
- Travis Lowdermilk. *User-centered Design: a Developer’s Guide to Building User-friendly Applications*. ” O’Reilly Media, Inc.”, 2013.
- Huawei Shen, Yuanhao Liu, Kaike Zhang, Qi Cao, and Xueqi Cheng. The rising safety concerns of deep recommender systems. *The Innovation*, 6(10), 2025.

- 
- Ben Shneiderman. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6):495–504, 2020.
- Ítallo Silva, Leandro Marinho, Alan Said, and Martijn C Willemsen. Leveraging chatgpt for automated human-centered explanations in recommender systems. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*, pages 597–608, 2024.
- Shoujin Wang, Xiuzhen Zhang, Yan Wang, and Francesco Ricci. Trustworthy recommender systems. *ACM Transactions on Intelligent Systems and Technology*, 15(4):1–20, 2024.
- Stanisław Woźniak, Jacek Duszenko, Jan Kocoń, and Przemysław Kazienko. Improving llm-based recommender systems with user-controllable profiles. In *Companion Proceedings of the ACM on Web Conference 2025*, pages 2102–2111, 2025.
- Yunfan Wu, Qi Cao, Shuchang Tao, Kaike Zhang, Fei Sun, and Huawei Shen. Accelerating the surrogate retraining for poisoning attacks against recommender systems. In *Proceedings of the 18th ACM Conference on Recommender Systems*, pages 701–711, 2024.
- Xin Xin, Liu Yang, Ziqi Zhao, Pengjie Ren, Zhumin Chen, Jun Ma, and Zhaochun Ren. On the effectiveness of unlearning in session-based recommendation. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, pages 855–863, 2024.
- Wujiang Xu, Yunxiao Shi, Zujie Liang, Xuying Ning, Kai Mei, Kun Wang, Xi Zhu, Min Xu, and Yongfeng Zhang. Instructagent: Building user controllable recommender via llm agent. *arXiv e-prints*, pages arXiv–2502, 2025.
- Kaike Zhang, Qi Cao, Fei Sun, Yunfan Wu, Shuchang Tao, Huawei Shen, and Xueqi Cheng. Robust recommender system: a survey and future directions. *ACM Computing Surveys*, 58(1): 1–38, 2025a.
- Kaike Zhang, Jiakai Tang, Du Su, Shuchang Liu, Julian McAuley, Lina Yao, Qi Cao, Yue Feng, and Fei Sun. The 2nd workshop on human-centered recommender systems. *arXiv preprint arXiv:2511.19979*, 2025b.
- Kaike Zhang, Yunfan Wu, Yougang Lyu, Du Su, Yingqiang Ge, Shuchang Liu, Qi Cao, Zhaochun Ren, and Fei Sun. The 1st workshop on human-centered recommender systems. In *Companion Proceedings of the ACM on Web Conference 2025*, pages 2088–2091, 2025c.