



GENNEXT: The Next Generation of IR and Recommender Systems with Language Agents, Generative Models, and Conversational AI

Yashar Deldjoo
deldjooy@acm.org
Polytechnic University of Bari
Bari, Italy

Scott Sanner
ssanner@mie.utoronto.ca
University of Toronto
Toronto, Canada

Enrico Palumbo
enricop@spotify.com
Spotify
Turin, Italy

Hugues Bouchard
hb@spotify.com
Spotify
Barcelona, Spain

Shuai Zhang
shuaizs@amazon.com
Amazon Web Services
Santa Clara, United States

Pablo Castells
pbl.castells@gmail.com
Autonomous University of Madrid
Madrid, Spain

Julian McAuley
jmcauley@ucsd.edu
UC San Diego
San Diego, USA

Abstract

We present **GENNEXT**, a workshop dedicated to exploring the integration of language agents, generative models, and conversational AI within information retrieval (IR) and recommender systems (RS). Building on the success of our recent RecSys'24 workshop, **GENNEXT** aims to advance discussions on the applications of *language agents* powered by Large Language Models (LLMs). The workshop will focus on enhancing interactivity between users and systems through multi-turn dialogues, improving creative content generation, advancing personalization, and enabling multifaceted, context-aware decision-making. For example, a language agent could respond to a query like “Suggest an eco-friendly food tour for a weekend in my city” by using a recommendation API to identify eateries specializing in sustainable or organic cuisine and a pollution API to ensure the selected routes have low air pollution levels.

GENNEXT will bring together leading researchers and practitioners through keynotes, paper presentations, and a panel discussion. We invite full papers, short papers, and extended abstracts covering theoretical advancements, practical applications, and evaluation strategies for generative technologies in IR and RS. The workshop will address key themes such as conversational adaptation, generative content creation, and agentic tool usage, while tackling challenges like bias, data privacy, and hallucination risks. Overall, our main ambition is to foster dialogue on creating ethical, sustainable, and innovative systems while addressing emerging opportunities and risks in modern IR and RS.

CCS Concepts

• **Information systems** → *Recommender systems*.



This work is licensed under a Creative Commons Attribution 4.0 International License. *SIGIR '25, Padua, Italy*

© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1592-1/2025/07
<https://doi.org/10.1145/3726302.3730369>

Keywords

Generative Models, Recommender Systems, Information Retrieval, Language Agents, LLMs, Conversational AI, Personalization, Ethical AI

ACM Reference Format:

Yashar Deldjoo, Scott Sanner, Enrico Palumbo, Hugues Bouchard, Shuai Zhang, Pablo Castells, and Julian McAuley. 2025. GENNEXT: The Next Generation of IR and Recommender Systems with Language Agents, Generative Models, and Conversational AI. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '25)*, July 13–18, 2025, Padua, Italy. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3726302.3730369>

1 Motivation and Relevance to SIGIR

Modern Information Retrieval (IR) and Recommender Systems (RS) are experiencing a profound change with the advent of Large Language Models (LLMs) [4, 5]. Traditional algorithms often rely on static features or past user-item interactions, whereas *language-agent* systems dynamically integrate world knowledge, language understanding, reasoning, and planning abilities to improve and expand the capabilities of IR and RS in a tangible manner [7, 16–18].

At a high level, language agents typically include an LLM component and a set of tools that the LLM can use. The LLM component can reason about complex queries, model nuanced user taste profiles, generate text for explaining results, and ask follow-ups, and refines queries in conversational setups. Crucially, the LLM also decides which tool(s) to invoke (e.g., traditional recommender and search systems, general purpose APIs) for improved accuracy and relevance. For instance, given a complex request such as “Find a good jacket for me to buy for my trip to New York next week” issues to a language agent, the LLM component may decide to invoke first a Weather API to check the weather forecasts in New York City in the following week, then reason on what type of jacket is most appropriate to that weather, and finally invoke a recommendation API to suggest a curated set of products. This pipeline shows how LLM-driven tool usage can support users through more personalized and context-sensitive recommendations.

Beyond improving retrieval and recommendation accuracy, language agents can also drive richer *conversational interactions* that adapt to user feedback. For example, a language agent might provide text-based feedback on why a particular movie was appealing, enabling the recommender engine to fine-tune its model parameters accordingly. This refined feedback loop can boost both user satisfaction—by suggesting more relevant content—and engagement time, as users spend longer interacting with recommendations tailored to their evolving preferences. More broadly, by clarifying user intents on the fly, such agents transform static one-shot recommendations into dynamic, multi-turn dialogues.

Additionally, LLM-based strategies enable the *generation* of novel or repurposed *content*. For instance, they may create virtual try-on experiences in online fashion platforms or stitch together personalized playlists that intuitively match a user's listening history. These advanced generation capabilities open new opportunities for diversifying item catalogs and providing richer user experiences beyond static, pre-existing content repositories.

Despite these benefits, deploying language agents in IR and RecSys poses challenges [9]. The risk of hallucinations—where agents produce inaccurate or misleading content—remains a concern, along with biases inherited from large-scale training data. An open research problem is also how to enable language agents to effectively and efficiently learn from environment rewards such as user interactions. Our proposed **GENNEXT@SIGIR'25** workshop focuses on identifying the new research frontiers, opportunities, and risks as language agents and generative models become central to IR and recommender ecosystems.

Why Appropriate for SIGIR? SIGIR has a long tradition of cutting-edge research in information retrieval, recommendation, and related areas. The official *Call for Papers* for SIGIR 2025¹ highlights several categories directly resonating with our workshop scope, including *Machine Learning for IR, Generative IR, Conversational IR and Intelligent Agents, Recommender Systems, and Multi-Modal IR*. Our workshop focuses on advanced LLM-driven frameworks, real-time user feedback loops, and cross-modal content generation—topics that align well with SIGIR's calls on generative retrieval, large language models, multi-turn interactions, and beyond. By uniting domain experts and practitioners in a dedicated forum, we aim to complement the main SIGIR 2025 conference topics while spotlighting specialized issues (e.g., multi-turn dialogues, agent decisions, tool integration) that are not typically covered in-depth in the main program.

Interest to the IR Community: We foresee wide interest in bridging IR, conversational AI, generative modeling, language agents, and recommender systems. The workshop's interactive format and focus on real-world deployment challenges (bias, evaluation, etc.) will draw both academic and industrial participants.

2 Theme and Purpose

The workshop's overarching goal is to explore the **next generation** of IR and recommendation systems, where LLMs and language agents serve as core components. We focus on:

- (1) **Conversational Adaptation and Multi-turn Interaction:** How can we design agents that handle complex mixed initiative sequential interactions with users, refine results on the fly, and adapt to shifting contexts?
- (2) **Agentic Tool Usage:** What APIs or specialized modules (recommender, QA system, knowledge graph, etc.) are most beneficial to integrate with LLMs and how? How do agents learn to select and call these tools effectively?
- (3) **Generative Content Creation:** Methods for generating or repurposing content, such as novel items or item compositions, creative writing, or immersive experiences, and how such capabilities intersect with classical IR/RecSys goals.
- (4) **Evaluation, Bias, and Fairness:** New metrics for measuring agent-driven IR and RecSys performance, especially with multi-turn dialogues or generative outputs. Approaches to mitigate bias and ensure ethical deployment.

Workshop Scope

Note that, inspired by our RecSys'24 workshop on Risks, Opportunities, and Evaluation of Generative Models in Recommender Systems (ROEGEN@RecSys'24),^a this workshop welcomes broader generative model research as long as it relates to IR, recommender systems, and conversational AI. Topics of interest include LLM fine-tuning, in-context learning, evaluation methodologies, and addressing challenges such as bias, fairness, security, privacy of generative models, and model hallucinations. Additionally, we encourage submissions exploring multimodal integration, which combines diverse data sources (e.g., text, images, audio) in particular in the context of language agents to enhance personalization and user satisfaction.

^a<https://roegen-recsys2024.github.io/>

3 Workshop Format, Planned Activities, and Schedule

We propose a **full-day** workshop structured around a balance of *invited keynotes*, *paper presentations*, and highly *interactive sessions*:

- **Invited Keynotes:** We anticipate at least one keynote speaker from an industry leader in Generative AI and one from academia. One keynote will focus on fairness or ethical issues.
- **Paper Presentations:** Accepted papers (long, short, or extended abstracts) will be organized into thematic sessions for example, focusing on LLM-based retrieval, conversation management, generative content, etc.
- **Panel/Breakout Discussions:** Small-group interactions around selected topics (e.g., mitigating hallucination, multi-step user simulation, etc), will be organized by the end of the workshop in a plenary panel setting and/or multiple breakout groups.
- **Poster/Demo Session:** We encourage demos of agent-based IR or next-generation RecSys prototypes.

Tentative Schedule.

09:00–09:15 Opening + Intro

09:15–10:15 Keynote 1

¹<https://sigir2025.dei.unipd.it/call-full-papers.html>

10:15–10:30 Coffee Break
10:30–12:00 Paper Presentations (Session I)
12:00–13:30 Lunch
13:30–14:30 Keynote 2
14:30–15:30 Paper Presentations (Session II)
15:30–15:45 Coffee Break
15:45–16:45 Breakouts / Panel Discussion
16:45–17:30 Poster/Demo Session + Wrap-Up

3.1 Planned Interaction and Engagement

To differentiate from a “mini-conference,” we will:

- Use **live Q&A** and a moderated panel to address open problems (e.g., agent hallucination, user modeling).
- Host **small-group breakouts** on specialized subtopics (bias and fairness, data privacy, domain-adaptive LLMs).
- Encourage **demo lightning talks** to showcase new agentic IR or generative RecSys prototypes.

This format ensures participants can network, brainstorm collaboratively, and share hands-on experiences.

4 Distinction from Main Conference Topics

While SIGIR 2025 will feature work on IR foundations, neural ranking, and possibly some papers on LLM-based recommendation, our workshop provides:

- A dedicated forum for *language-agent frameworks* (choosing tools, chaining calls, complex reasoning).
- Emphasis on *generative content* and *multimodal synergy* (beyond text-only retrieval).
- A focus on *conversational interactivity* and real-time user feedback.
- Attention to risks such as *bias, privacy, and ethics* for agentic IR/RecSys.

This complements the main conference by uniting IR, RecSys, and broader NLP/LLM experts specifically around agent-based and generative use-cases.

5 Audience Reach

We anticipate attracting:

- **Academics** in IR, RecSys, or NLP bridging LLM-based approaches.
- **Industry practitioners** building next-generation search and recommendation services.
- **Researchers in fairness/ethics**, focusing on data bias, user trust, or explanation in agent-based IR systems.
- **Students / newcomers** seeking to gain exposure to advanced LLM-based IR or recommendation paradigms.

6 Special Requirements and Contingency Plans

Onsite Attendance: All organizers plan to attend in person. If unforeseen circumstances arise (e.g., travel issues), at least two of us will be onsite. We also have local collaborators who can help run sessions.

Technical Requirements: We only need standard projector/audio support, poster boards, and space for breakout sessions. If remote keynote involvement becomes necessary, we will coordinate with SIGIR organizers for streaming.

7 List of Organizers and Short Biographies

Yashar Deldjoo (Polytechnic University of Bari)

Assistant Professor specializing in recommender systems and trustworthy AI [2, 3, 6, 10–12]. Extensive research on multi-modal RS and new ventures into LLM-driven personalization. Associate Editor at ACM Computing Surveys.

Scott Sanner (University of Toronto)

Professor researching AI, sequential decision-making, and deep learning with information retrieval and recommendation applications [4, 8, 14]. Associate Editor for ACM TORS, MLJ, and JAIR and recipient of multiple Google Faculty Research Awards.

Enrico Palumbo (Spotify)

Research Scientist focusing on language technologies for search and recommendation [1, 13]. Recent emphasis on LLM-based approaches, synergy between dialogue modeling, retrieval, and personalization.

Hugues Bouchard (Spotify)

Senior Research Manager focusing on language technologies and generative models to advance search and recommendation at scale, and unlock new GenAI interactive and listening experiences.

Shuai Zhang (Amazon Web Services)

Applied Scientist specializing in foundation models for IR and graph ML [6, 15]. Contributed to Amazon Titan, OpenSearch, and novel neural retrieval strategies.

Pablo Castells (Autonomous University of Madrid / Amazon Scholar)

Professor with a strong background in IR and recommender systems. Interests in evaluation design, bias, novelty, diversity, and interactive recommendation.

Julian McAuley (UC San Diego)

Professor whose group pioneered neural approaches for personalization and text/image integration, with industry collaborations (Adobe, Pinterest, Etsy, Microsoft, Amazon).

8 Potential Program Committee and Selection Process

Program Committee: We will invite experts from academia and industry with specialized focus on:

- LLM-based IR, generative retrieval
- Advanced recommendation (conversational, multi-modal)
- Fairness, ethics, transparency
- Evaluation and user studies

Submission and Review: We will solicit papers via EasyChair. Each submission will receive at least two reviews based on relevance, technical merit, novelty, and clarity. Final decisions will balance reviewer feedback, diversity of topics, and potential for stimulating discussion.

9 Expected Target Audience and Advertising

We expect researchers, practitioners, and graduate students interested in bridging IR, recommendation, and LLM-based AI. We will promote through:

- SIGIR mailing lists and social media channels (Twitter, LinkedIn).
- Relevant labs and developer communities in IR/RecSys/NLP.

- Invitations to prior participants of similar workshops (Gen-IR@CIKM, LLMAI@KDD).

10 Related Workshops

Recent workshops such as **LLMAI@KDD**, **WSDM 2024's LLM for Society**, and **Gen-IR@CIKM** have explored large language models in various contexts. Our workshop is unique in its dedicated focus on *agentic frameworks* for IR and recommendations, including *multi-turn dialogues*, *tool orchestration*, and *generative content creation*—all with an eye toward evaluating and mitigating biases or hallucination risks.

11 Submission Types and Key Dates

Submission Types

We encourage five types of submissions to cover a broad range of research contributions and perspectives. Each type is designed to suit different stages of research and insights:

- **Full Papers:** Maximum 8 pages (excluding references), suitable for presenting comprehensive studies and detailed research findings.
- **Short Papers:** Maximum 4 pages (excluding references), ideal for sharing preliminary results or innovative ideas.
- **Extended Abstracts:** 2–3 pages (excluding references), focusing on work-in-progress, position statements, or early-stage concepts.
- **Perspective Papers:** Focused discussions on future directions, challenges, or thought-provoking ideas in generative and agentic IR/RS. Maximum 8 pages.
- **Mini-Reviews:** Concise (mini) reviews that critically study current trends, gaps, and opportunities in the field. Maximum 8 pages.

Camera-ready versions of accepted papers will be included in the workshop proceedings (likely CEUR-WS). Submissions must follow ACM's two-column sigconf format and be anonymized for review.

Key Dates

The following timeline highlights important milestones for the workshop. These dates are chosen to align with the SIGIR 2025 schedule:

- **Submission Deadline:** April 23, 2025
Researchers are encouraged to submit their work by this date to ensure it is considered for review.
- **Notification:** May 21, 2025
Accepted papers will be notified, allowing sufficient time for authors to revise their work.
- **Camera-Ready Deadline:** (Tentative) June 3, 2025
Final versions of accepted papers must be submitted for inclusion in the proceedings.
- **Workshop Date:** July 17, 2025
The workshop will take place during the SIGIR 2025 conference in Padua, Italy.

Acknowledgments

We thank the SIGIR 2025 organizers and community for providing an excellent venue to discuss cutting-edge topics in IR and Recommender Systems. Special thanks to all who gave input on the workshop scope and structure.

References

- [1] Andreas Damianou, Francesco Fabbri, Paul Gigoli, Marco De Nadai, Alice Wang, Enrico Palumbo, and Mounia Lalmas. 2024. Towards graph foundation models for personalization. In *Companion Proceedings of the ACM Web Conference 2024*. 1798–1802.
- [2] Yashar Deldjoo. 2024. Understanding Biases in ChatGPT-based Recommender Systems: Provider Fairness, Temporal stability, and Recency. *ACM Transactions on Recommender Systems* (2024).
- [3] Yashar Deldjoo and Tommaso Di Noia. 2024. CFairLLM: Consumer Fairness Evaluation in Large-Language Model Recommender System. *ACM Transactions on Intelligent Systems and Technology* (2024).
- [4] Yashar Deldjoo, Zhankui He, Julian McAuley, Anton Korikov, Scott Sanner, Arnau Ramisa, René Vidal, Maheswaran Sathiamoorthy, Atoosa Kasrizadeh, and Silvia Milano. 2024. A Review of Modern Recommender Systems using Generative Models (Gen-RecSys). In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 6448–6458.
- [5] Yashar Deldjoo, Zhankui He, Julian McAuley, Anton Korikov, Scott Sanner, Arnau Ramisa, René Vidal, Maheswaran Sathiamoorthy, Atoosa Kasrizadeh, Silvia Milano, et al. 2024. Recommendation with generative models. *arXiv preprint arXiv:2409.15173* (2024).
- [6] Yashar Deldjoo, Nikhil Mehta, Maheswaran Sathiamoorthy, Shuai Zhang, Pablo Castells, and Julian McAuley. 2025. Toward Holistic Evaluation of Recommender Systems Powered by Generative Models. *SIGIR'25* (2025).
- [7] Luke Friedman, Sameer Ahuja, David Allen, Zhenning Tan, Hakim Sidahmed, Changbo Long, Jun Xie, Gabriel Schubiner, Ajay Patel, Harsh Lara, et al. 2023. Leveraging large language models in conversational recommender systems. *arXiv preprint arXiv:2305.07961* (2023).
- [8] Sara Kemper, Justin Cui, Kai Dicarlantonio, Kathy Lin, Danjie Tang, Anton Korikov, and Scott Sanner. 2024. Retrieval-Augmented Conversational Recommendation with Prompt-based Semi-Structured Natural Language State Tracking. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2786–2790.
- [9] Reza Yousefi Maragheh, Yashar Deldjoo, Chi Wang, Jason Cho, and Derek Cheng. 2025. Multi Agentic Recommender Systems: Foundations, Design Patterns, and E-Commerce Applications (An Industrial Tutorial). In *RecSys'25*.
- [10] Fatemeh Nazary, Yashar Deldjoo, and Tommaso di Noia. 2025. Poison-RAG: Adversarial Data Poisoning Attacks on Retrieval-Augmented Generation in Recommender Systems. *ECIR'25* (2025).
- [11] Fatemeh Nazary, Yashar Deldjoo, Tommaso Di Noia, and Eugenio Di Sciascio. 2025. Stealthy LLM-Driven Data Poisoning Attacks Against Embedding-Based Retrieval-Augmented Recommender Systems. *arXiv preprint arXiv:2505.05196* (2025).
- [12] Fatemeh Nazary, Yashar Deldjoo, Tommaso Di Noia, and Eugenio Di Sciascio. 2025. XAI4LLM. Let Machine Learning Models and LLMs Collaborate for Enhanced In-Context Learning in Healthcare. *arXiv preprint arXiv:2405.06270* (2025).
- [13] Gustavo Penha, Ali Vardasbi, Enrico Palumbo, Marco De Nadai, and Hugues Bouchard. 2024. Bridging Search and Recommendation in Generative Retrieval: Does One Task Help the Other?. In *Proceedings of the 18th ACM Conference on Recommender Systems*. 340–349.
- [14] Suvash Sedhain, Aditya Krishna Menon, Scott Sanner, and Lexing Xie. 2015. Autorec: Autoencoders meet collaborative filtering. In *Proceedings of the 24th international conference on World Wide Web*. 111–112.
- [15] Bingqing Song, Boran Han, Shuai Zhang, Hao Wang, Haoyang Fang, Bonan Min, Yuyang Wang, and Mingyi Hong. 2025. Effectively Steer LLM To Follow Preference via Building Confident Directions. *arXiv preprint arXiv:2503.02989* (2025).
- [16] Lei Wang, Jingsen Zhang, Hao Yang, Zhiyuan Chen, Jiakai Tang, Zeyu Zhang, Xu Chen, Yankai Lin, Ruihua Song, Wayne Xin Zhao, et al. 2023. User behavior simulation with large language model based agents. *arXiv preprint arXiv:2306.02552* (2023).
- [17] Zhefan Wang, Yuanqing Yu, Wendi Zheng, Weizhi Ma, and Min Zhang. 2024. MACRec: A Multi-Agent Collaboration Framework for Recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2760–2764.
- [18] An Zhang, Yuxin Chen, Leheng Sheng, 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.