








# Workshop on Conversational Search for Complex Information Needs

Roxana Petcu<sup>1</sup>(✉) , Mert Yazan<sup>2,3</sup> , Mohanna Hoveyda<sup>4</sup> , Jirui Qi<sup>5</sup> ,  
and Maarten de Rijke<sup>1</sup> 

<sup>1</sup> University of Amsterdam, Amsterdam, The Netherlands  
r.m.petcu@uva.nl

<sup>2</sup> Amsterdam University of Applied Science, Amsterdam, The Netherlands

<sup>3</sup> Leiden University, Leiden, The Netherlands

<sup>4</sup> Radboud University, Nijmegen, The Netherlands

<sup>5</sup> University of Groningen, Groningen, The Netherlands

**Abstract.** We present a workshop proposal on *Conversational AI and Search for Complex Information Needs*. Recent advances in conversational systems have expanded research beyond generative dialogue and factoid question-answering tasks, towards addressing complex information needs and improving user experience by providing adaptive answers to users based on style and personalization. In conversational search, user queries first go through intermediary steps (e.g., with agentic flows) before providing an answer that fits the user's expectations. As a result, research has shifted its focus to conversations that support exploratory, multi-step information needs, and to challenges such as long-form answer generation, personalization, the orchestration of agentic, and societal considerations, which remain largely open areas of research. This workshop is open to researchers and practitioners from the information retrieval, natural language processing, and human-computer interaction backgrounds to exchange results and prospective ideas around conversational AI by discussing the impacts on users and the need for adaptivity. The goal and expected outcome of the workshop is to bring together diverse perspectives, foster mutual learning and collaboration, and build better systems in the future. The program will combine poster presentations, interactive breakout sessions, and invited speakers.

**Keywords:** Conversational AI · Information Retrieval · Complex Information Needs · Personalization · Societal Impact

## 1 Motivation

Conversational AI fosters a powerful environment in which people interact with information. Throughout this proposal, we use a few key terms: **Conversational search** refers to a multi-turn dialogue setting in which the user's information need evolves: At each turn, the system must interpret the underlying intent and generate an answer that effectively addresses it; **Conversational AI** takes a step

back and refers to broader methods and pipelines such as Retrieval Augmented Generation (RAG) or agentic orchestration, to address user information needs as precisely as possible; **Information need** refers to the underlying question of the user that might be expressed explicitly or not; **Complex Information Need** refers to a user request that often can be open-ended, and often requires multiple sources of information, reasoning or planning to answer [2,9,21].

Unlike traditional search, where users must write an explicit query that mirrors their information need and navigate through ranked lists from which they have to extract information themselves, conversational systems allow for all these steps through a natural, adaptive dialogue setting. It is the task of the conversational system to rephrase the user query, extract information from relevant sources, and combine them in a comprehensive answer. As users gain more experience with conversational systems, new expectations appear. Agents are increasingly expected to be experts in specific domains, such as education, healthcare, or finance. They might also be expected to keep track of user-specific information, such as dietary restrictions or movie preferences, so that the user does not have to state them each time when relevant information is needed.

To truly personalize the interactions, it is essential to understand how a user expresses themselves to understand their true expectations and intentions, and to avoid any biases due to personal backgrounds. At the same time, it becomes essential for conversational systems to become reliable and robust, as users tend to spend less time verifying information and are prone to use them more often than before, and even in sensitive domains. This shows the need for conversational systems to be adaptive, to address the potential problems that might arise given the increasing expectations.

**Relevance to ECIR:** The European Conference on Information Retrieval (ECIR) has been at the forefront of Information Retrieval open research, supporting advances in theoretical and applied breakthroughs. Meanwhile, recent topics such as Conversational AI, Retrieval Augmented Generation (RAG), and evaluation have positioned themselves as a collaborative fusion between IR and Natural Language Processing (NLP). This workshop aligns with ECIR’s focus on cutting-edge research in Information Retrieval by aiming to find novel works on complex user scenarios, personalization, and user modeling, promoting responsible and explainable systems, and bringing together the science community.

## 2 Goals

The goal of this workshop is to bring together researchers, practitioners, and industry partners to advance the study of conversational AI and search for complex information needs. Beyond simply presenting recent results, the workshop is intended to foster dialogue across disciplines, encourage the exchange of ideas, put the focus on user needs and expectations, and inspire new directions. By convening experts from information retrieval, natural language processing, and human-computer interaction, we aim to surface shared challenges, highlight

opportunities for collaborations, and stimulate exploration of understudied but impactful research areas.

**Vision.** We envision this workshop as more than a sequence of talks and posters. It is expected to be an opportunity to bring together a diverse community to reflect on the state of conversational search and identify key challenges for the future. Through invited talks, presentations, and poster sessions, participants will gain a broad view of ongoing work, emerging approaches, and open research questions. This shared exposure will help create a clearer sense of where the field stands today and which directions hold the greatest promise. We aim to spark thoughtful discussion about the limitations of current systems—including issues of trustworthiness, bias, and user adaptation—and to inspire research that makes conversational systems more robust, transparent, and genuinely helpful. By expanding the scope to other areas such as human-computer interaction, we wish to initiate a new outlook where conversational system needs are aligned with user needs and expectations.

**Outcomes.** We expect this workshop to generate both immediate and lasting benefits for the community. In the short term, we aim for the participants to not only have a chance to discuss ideas and challenges they face, but also to meet colleagues who work on complementary topics in other fields. Therefore, this workshop is aimed at building a community for researchers who contribute to the design of conversational search that is adaptive and addresses complex information needs across various fields. In the longer term, we anticipate that discussions initiated at the workshop will contribute to the design of shared tasks, benchmarks, and reproducibility efforts that can accelerate progress and support more developments in conversational search for complex information needs. Ultimately, we aspire to strengthen a community dedicated to developing conversational systems that are not only technically capable but also human-centered, trustworthy, and responsive to the complex, evolving needs of users.

### 3 Structure

The workshop will be organized as a balanced program of presentations of new research and invited talks from experts in the field. In this manner, we encourage an interactive learning environment for everyone involved in the workshop. In this section, we outline specifics of the workshop, such as the planned duration of the event, types of submissions we encourage, style of presentation, tentative schedule, and organizing team.

*Duration.* We propose a **half-day workshop** of 3.5 h plus breaks.

*Type of Papers.* We encourage contributions from both mature and early-stage work. We plan to accept technical papers of **4 to 8 pages**. We also invite **position papers** up to 4 pages. The workshop will be non-archival.

*Workshop Subtopics.* We propose multiple aspects of Conversational AI and Search, and welcome submissions in any of the following areas:

1. **Complex Information Needs:** User queries that go beyond factoid retrieval have been shown to pose significant challenges for retrieval systems [4, 15]. These queries require reasoning, synthesis of multiple pieces of information [19], or planning to fill knowledge gaps [7]. Such queries may involve implicit logical operations (e.g., negation, disjunction, conjunction) [12, 23], multi-step dependencies [6], topical complexity [8], or require information that lies outside the model’s training data [13, 18]. In this regard, we welcome contributions that propose (i) *evaluation frameworks or datasets* to effectively expose the limitations of current retrieval and RAG systems, including analyses of how learned representations encode logical structures that leads to performance failure, (ii) *models or frameworks* that more effectively handle complex information needs.
2. **Data Generation:** Methods for creating new benchmark datasets that capture complex conversational search needs [3]. This includes approaches for generating new datasets, augmenting existing datasets, and evaluating the created resources in comparison with existing benchmarks.
3. **Orchestration:** Modern information systems increasingly rely on coordinating multiple models, tools, and retrieval components [10]. Orchestration refers to adaptive and agentic approaches that dynamically decide how these resources should be combined to best address a user’s request [5]. This involves selecting answering strategies that enable adapting system behavior to the complexity of users’ queries. We thereby encourage contributions on (i) *orchestration mechanisms*, (ii) *optimization approaches*, and (iii) *evaluation under dynamic conditions*.
4. **Personalization and Adaptivity:** Personalization encompasses many areas, including personalized text generation, response adaptation to diverse user backgrounds, and context tracking. Personalized text generation refers to generating text in the style of an individual, and is especially relevant for query rewriting in the context of conversational systems, and helps the system to adapt to the user’s style [22]. In this workshop, we approach the personalization problem as an adaptivity problem, and tackle the problem of relating to users from diverse backgrounds, given the inherent biases and stereotypes of LLMs [1].
5. **User Experience:** Users report higher satisfaction with conversational search compared to traditional search engines, due to exercising less effort [11]. This, however, leads to new issues such as overreliance, where users blindly trust the answer they get from conversational interfaces without noticing that it is wrong [17]. Therefore, a new paradigm has emerged where users actively interact with both types of interfaces to balance their weak points [14]. Given this, there is a lack of studies understanding how people approach this new paradigm, and how their expectations and needs have changed with it.
6. **Bias, Fairness, Interpretability:** Responsible systems wrt. ethical and societal considerations, bias mitigation, inclusivity, and transparency [20].

Including works on mitigating misinformation and applications in sensitive domains.

7. **Evaluation:** Evaluation remains a key challenge for conversational search, as current metrics rely on pre-defined gold answers, failing to assess dimensions of the generated answer that go beyond mere overlap with some reference answer. Traditional relevance measures are insufficient for assessing long-form answers, faithfulness of reasoning steps, context tracking, or user-centered aspects such as safety, bias, and personalization [16]. We call for contributions addressing these gaps with new metrics and evaluation approaches.

*Types of Presentations.* We propose two types of presentations, i.e., **oral presentations** for the keynotes, and **short pitches and posters** for submitted and accepted papers.

*Planned Activities.* Rather than organizing a mini-conference, we plan to hold a highly interactive event, with keynote presentations to set the scene, poster presentations to develop a shared background, and a panel to flesh out key challenges and directions.

#### *Tentative Schedule*

09:00–09:10: Opening remarks and introduction  
 09:10–09:30: Keynote 1  
 09:30–10:15: Panel  
 10:15–10:30: Coffee break  
 10:30–10:40: One-minute poster pitches  
 10:40–11:40: Poster session  
 11:40–12:00: Keynote 2  
 12:00–12:30: Wrap-up and closing remarks

*Resources Needed.* To support our proposed workshop, we need to facilitate a combination of keynotes, short pitches, posters, and a panel. For these, we will need a projector for keynote presentations, with two to three microphones to accommodate the speaker, mediator, and potential questions from the audience. Moreover, for the poster presentation, we will need several poster boards. A system that allows for easy connection on the audio-video support, and internet connection is necessary. Finally, for the panel, we need to be able to organize the room in a suitable way.

## 4 Community

### 4.1 Intended Audience

The workshop is intended for researchers and practitioners working in conversational AI, information retrieval, natural language processing, or human-computer

interaction. We expect a particular interest in one of the subtopics presented in this proposal and hope that bringing research from all these areas together will facilitate a global understanding of recent advances and remaining research gaps in Conversation AI and Search. We expect between 20–30 participants. Participants are welcomed from academia, including master’s students, PhDs, and other senior roles, alongside contributors from industry.

## 4.2 Organization Team

The workshop will be organized by researchers from multiple research labs with expertise in conversational search, personalization, fairness, and evaluation. The organizing team brings a unique perspective on conversational AI, ensuring proficiency in all research aspects of the workshop.

- Roxana Petcu (University of Amsterdam) - PhD researcher focusing on complex user needs, complex queries, topological phenomena in language, query decomposition, RAG, and data generation.
- Mert Yazan (Amsterdam University of Applied Sciences) - PhD researcher with expertise in personalization, user modeling, user studies, and societal impact of conversational agents.
- Mohanna Hoveyda (Radboud University) - PhD researcher working on methods for complex information needs, with a focus on probabilistic reasoning frameworks and neuro-symbolic systems.
- Jirui Qi (University of Groningen) - PhD researcher on evaluation benchmarks, knowledge consistency, and model explainability.
- Maarten de Rijke (University of Amsterdam) - Distinguished university professor, working on trustworthy information retrieval technology

## 4.3 Program Committee Members

For the program committee members, we will invite academic and industrial researchers with expertise in the above-presented subtopics from around the world:

- Krisztian Balog (U. Stavanger and Google DeepMind)
- Samarth Bhargav (Cohere)
- Yixing Fan (Chinese Academy of Sciences)
- Faegheh Hasibi (Radboud U. Nijmegen)
- Dawn Laurie (Johns Hopkins U.)
- Dan Li (Elsevier)
- Chuan Meng (U. Edinburgh)
- Alessandro Moschitti (Amazon)
- Pengjie Ren (Shandong U.)
- Svitlana Vakulenko (Vienna U. of Economics and Business)
- Hamed Zamani (U. Massachusetts)
- Guglielmo Faggioli (University of Padova)

#### 4.4 Invited Speakers and Panel Members

We invited two keynote speakers, more precisely Mohammad Aliannejadi (University of Amsterdam), and Cristina Ioana Muntean (Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"). The panel will be held to flourish discussions starting from and around the work of Yougang Lyu.

**Disclosure of Interests.** The authors have no relevant financial or non-financial interests to disclose.

#### References

1. Plaza-del Arco, F.M., Curry, A.C., Paoli, S., Cercas Curry, A., Hovy, D.: Divine LLaMAs: Bias, stereotypes, stigmatization, and emotion representation of religion in large language models. In: Al-Onaizan, Y., Bansal, M., Chen, Y.N. (eds.) Findings of the Association for Computational Linguistics: EMNLP 2024, pp. 4346–4366. Association for Computational Linguistics, Miami, Florida, USA (2024). <https://doi.org/10.18653/v1/2024.findings-emnlp.251>
2. Asai, A., Hashimoto, K., Hajishirzi, H., Socher, R., Xiong, C.: Learning to retrieve reasoning paths over wikipedia graph for question answering (2020). <https://arxiv.org/abs/1911.10470>
3. Askari, A., et al.: Self-seeding and multi-intent self-instructing LLMs for generating intent-aware information-seeking dialogs (2025). <https://arxiv.org/abs/2402.11633>
4. van den Elsen, C., Barkhof, F., Nijdam, T., Lupart, S., Aliannejadi, M.: Reproducing NevIR: negation in neural information retrieval. In: Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '25 (2025)
5. Hoveyda, M., Oosterhuis, H., de Vries, A.P., de Rijke, M., Hasibi, F.: Adaptive orchestration of modular generative information access systems. In: Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (2025)
6. Jeong, S., Baek, J., Cho, S., Hwang, S.J., Park, J.: Adaptive-RAG: learning to adapt retrieval-augmented large language models through question complexity. In: Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (2024)
7. Jin, B., et al.: Search-R1: Training LLMs to reason and leverage search engines with reinforcement learning. In: Second Conference on Language Modeling (2025)
8. Killingback, J., Zamani, H.: Benchmarking information retrieval models on complex retrieval tasks. arXiv preprint [arXiv:2509.07253](https://arxiv.org/abs/2509.07253) (2025)
9. Kim, T., et al.: Learning to explore and select for coverage-conditioned retrieval-augmented generation (2025). <https://arxiv.org/abs/2407.01158>
10. Li, X., et al.: Chain-of-knowledge: grounding large language models via dynamic knowledge adapting over heterogeneous sources. In: The Twelfth International Conference on Learning Representations (2024)
11. Liang, Y., Wu, Z., Zhang, F., Song, D., Huang, H.: How users interact with generative information retrieval systems: a study of user behavior and search experience. In: Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 634–644. SIGIR '25, Association

- for Computing Machinery, New York, NY, USA (2025). <https://doi.org/10.1145/3726302.3729998>
12. Malaviya, C., Shaw, P., Chang, M.W., Lee, K., Toutanova, K.: QUEST: a retrieval dataset of entity-seeking queries with implicit set operations. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (2023)
  13. Mallen, A., Asai, A., Zhong, V., Das, R., Khashabi, D., Hajishirzi, H.: When not to trust language models: Investigating effectiveness of parametric and non-parametric memories. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (2023)
  14. Menon, D., Shilpa, K.: “Chatting with ChatGPT”: analyzing the factors influencing users’ intention to use the OpenAI’s ChatGPT using the UTAUT model. *Heliyon* 9(11), e20962 (2023). <https://doi.org/10.1016/j.heliyon.2023.e20962>
  15. Petcu, R., Bhargav, S., de Rijke, M., Kanoulas, E.: A comprehensive taxonomy of negation for NLP and neural retrievers. arXiv preprint [arXiv:2507.22337](https://arxiv.org/abs/2507.22337) (2025)
  16. Salemi, A., Killingback, J., Zamani, H.: ExPerT: Effective and explainable evaluation of personalized long-form text generation. In: Findings of the Association for Computational Linguistics: ACL 2025 (2025)
  17. Spatharioti, S.E., Rothschild, D., Goldstein, D.G., Hofman, J.M.: Effects of LLM-based search on decision making: speed, accuracy, and overreliance. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. CHI ’25, Association for Computing Machinery, New York, NY, USA (2025). <https://doi.org/10.1145/3706598.3714082>
  18. Su, W., Ai, Q., Zhan, J., Dong, Q., Liu, Y.: Dynamic and parametric retrieval-augmented generation. In: Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval. Association for Computing Machinery (2025)
  19. Trivedi, H., Balasubramanian, N., Khot, T., Sabharwal, A.: Interleaving retrieval with chain-of-thought reasoning for knowledge-intensive multi-step questions. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (2023)
  20. Wallat, J., Heuss, M., Rijke, M.d., Anand, A.: Correctness is not faithfulness in retrieval augmented generation attributions. In: Proceedings of the 2025 International ACM SIGIR Conference on Innovative Concepts and Theories in Information Retrieval (ICTIR), pp. 22–32. ICTIR ’25, Association for Computing Machinery, New York, NY, USA (2025). <https://doi.org/10.1145/3731120.3744592>
  21. Yao, S., et al.: React: Synergizing reasoning and acting in language models (2023). <https://arxiv.org/abs/2210.03629>
  22. Yazan, M., Verberne, S., Situmeang, F.: Improving RAG for personalization with author features and contrastive examples. In: Advances in Information Retrieval: 47th European Conference on Information Retrieval, ECIR 2025, Lucca, Italy, April 6–10, 2025, Proceedings, Part III, pp. 408–416. Springer-Verlag, Berlin, Heidelberg (2025). [https://doi.org/10.1007/978-3-031-88714-7\\_40](https://doi.org/10.1007/978-3-031-88714-7_40)
  23. Zhang, W., et al.: ExcluIR: Exclusionary neural information retrieval. In: Proceedings of the AAAI Conference on Artificial Intelligence (2025)