ProActLLM: Proactive Conversational Information Seeking with Large Language Models

Shubham Chatterjee*
Missouri University of Science and Technology
Rolla, USA
shubham.chatterjee@mst.edu

Xi Wang*
University of Sheffield
Sheffield, UK
xi.wang@sheffield.ac.uk

Shuo Zhang
Bloomberg
London, UK
szhang611@bloomberg.net

Sajad Ebrahimi
University of Guelph
Guelph, Canada
sebrah05@uoguelph.ca

Zhaochun Ren
Leiden University
Leiden, Netherlands
z.ren@liacs.leidenuniv.nl

Debasis Ganguly
University of Glasgow
Glasgow, UK
debasis.ganguly@glasgow.ac.uk

Gareth Jones
Dublin City University
Dublin, Ireland
gareth.jones@dcu.ie

Emine Yilmaz
University College London
London, UK
emine.yilmaz@ucl.ac.uk

Hamed Zamani
University of Massachusetts
Amherst, USA
zamani@cs.umass.edu

Abstract

Large Language Models (LLMs) have transformed information access by enabling human-like text understanding and generation. This workshop explores the next step for conversational AI: building proactive information-seeking assistants that go beyond reactive question answering. We aim to investigate how LLMs can anticipate user needs, model complex context, support mixed-initiative interactions, integrate retrieval and external tools, personalize responses, adapt through feedback, and ensure fairness, transparency, and cognitive grounding. Bringing together experts from NLP, IR, HCI, and cognitive science, the workshop will serve as a timely forum for advancing intelligent, proactive dialogue systems. It will also foster interdisciplinary collaboration.

ACM Reference Format:
Shubham Chatterjee, Xi Wang, Shuo Zhang, Sajad Ebrahimi, Zhaochun Ren, Debasis Ganguly, Gareth Jones, Emine Yilmaz, and Hamed Zamani. 2025. ProActLLM: Proactive Conversational Information Seeking with Large Language Models. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management (CIKM '25), November 10–14, 2025, Seoul, Republic of Korea.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3746252.3761593

1 Workshop Motivation

Large Language Models (LLMs) have recently demonstrated strong capabilities in understanding and generating human-like text, thereby transforming the way people access information. Despite these advances, many current information access strategies remain reactive [8]: users must formulate explicit queries, to express their intent, and systems retrieve information in response. A growing body of work now aims to shift this paradigm towards proactive and interactive information seeking [1, 4, 6].

With increased availability of data and the identification of key research challenges - such as user intent anticipation [5], context modelling [3], proactive decision-making [7], and the use of conversational interaction as a vehicle for proactivity [2] - this workshop focuses on a timely and important goal for conversational AI. The goal is moving beyond reactive question answering toward proactive assistants that anticipate user needs and deliver relevant information without requiring explicit queries.

The workshop aims to explore how LLMs can support improved information-seeking experiences through systems that model complex and dynamic contexts, anticipate evolving information needs, support mixed-initiative dialogues, integrate external tools and retrieval, generate personalised responses, adapt through feedback, and maintain fairness and transparency. We also seek to ground these systems in cognitive and pragmatic principles to support their deployment in real-world domains.

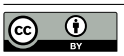
Given the recent progress in LLM-driven dialogue capabilities, now is an ideal time to bring together researchers and practitioners from natural language processing, information retrieval, human-computer interaction, and cognitive science. This workshop will provide a dedicated venue to share the latest findings, exchange ideas, and build collaborations in developing proactive conversational agents that can take initiative and improve information-seeking outcomes.

2 Workshop Theme and Topics

We invite submissions on a wide range of topics related to proactive conversational information seeking. Relevant topics include, but are not limited to:

(1) **Proactive question asking and clarification**: Techniques for an agent to formulate clarifying questions or prompts to better understand user intent.

*Both authors contributed equally to this research.



CIKM '25, Seoul, Republic of Korea

ACM ISBN 979-8-4007-2040-6/2025/11
https://doi.org/10.1145/3746252.3761593

(2) **Mixed-initiative dialogue management**: Models for balancing system initiative and user control in information-seeking conversations.
(3) **LLMs for conversational information access**: Leveraging LLMs to perform retrieval-augmented generation, knowledge grounding, and dynamic topic suggestion during dialogue.
(4) **User modelling and context understanding**: Inferring user goals, knowledge, and preferences throughout a conversation to drive proactive system actions.
(5) **Evaluation metrics and methodologies**: Developing metrics, benchmarks, and user-centric evaluation protocols for proactive conversational systems (e.g., measuring user satisfaction or engagement)
(6) **Domain-specific applications**: Proactive conversational assistants in domains such as healthcare, education (e.g., tutoring systems), customer support, enterprise search, or personal digital assistants.
(7) **Ethical and user experience considerations**: Ensuring proactive behaviour is helpful and not intrusive; handling user privacy, fairness for diverse users, and mitigating misinformation or biases in proactive responses.
(8) **Cognitive load management**: Designing proactive systems that consider users' cognitive capacity and attention limitations, timing interventions to minimize information overload while maximizing utility.
(9) **Memory and forgetting models**: Incorporating human-like memory mechanisms that prioritize, decay, and reconstruct information to inform proactive suggestions based on what users are likely to remember or forget.
(10) **Metacognitive awareness**: Building systems that can recognize when users are struggling with information processing, experiencing confusion, or need different types of cognitive support.
(11) **Expertise adaptation**: Dynamically adjusting proactive behavior based on users' domain expertise levels, from novice scaffolding to expert-level anticipatory support.
(12) **Multi-modal proactive assistance**: Integrating visual, audio, and textual cues to anticipate user needs across different interaction modalities and contexts.
(13) **Tool orchestration and API integration**: Seamlessly coordinating multiple external tools, databases, and services to fulfill complex, multi-step information needs.
(14) **Explainability and transparency**: Making proactive system decisions interpretable, and allowing users to understand why certain suggestions were made.
(15) **Cultural and linguistic adaptation**: Developing proactive systems that are sensitive to diverse cultural contexts and communication styles.

3 Workshop Objectives, Goals, and Expected Outcomes

ProActLLM addresses a timely need to advance conversational agents from reactive systems to proactive assistants that anticipate and support user needs. While traditional dialogue systems rely on explicit queries, recent progress in LLMs has opened new opportunities for anticipatory, mixed-initiative interaction. However, there is currently no dedicated venue that unifies this growing body of work. The workshop has five primary goals:

(1) to define and formalize the concept of proactivity in conversational systems, moving beyond conventional question-answering paradigms.
(2) to investigate how LLMs can model user intent, preferences, and latent goals over time in order to take initiative in a transparent and helpful way.
(3) to facilitate community discussion around evaluating proactive behavior, including metrics for user satisfaction, initiative quality, and long-term effectiveness.
(4) to bring together researchers from information retrieval, natural language processing, human-computer interaction, and cognitive science to examine the theoretical and practical foundations of proactive dialogue.
(5) to share insights from real-world deployments in domains such as education and healthcare, where proactive systems have the potential to significantly improve outcomes.

The expected outcomes of the workshop include:

(1) a collection of peer-reviewed papers capturing current research on proactive conversational systems;
(2) shared evaluation criteria and design guidelines grounded in user needs;
(3) stronger community connections across CIKM and related venues such as SIGIR, ACL, and CHI;
(4) open-access proceedings, invited talks, and continued collaboration through mailing lists and shared repositories

Ultimately, the workshop aims to define and advance an emerging subfield at the intersection of proactivity and conversational AI, supporting both foundational research and applied development.

4 Workshop Format and Target Audience

Workshop Length. ProActLLM is planned as a full-day workshop.

Workshop Program Format. ProActLLM will be a primarily in-person workshop with at least Shubham and Xi attending in person. The programme will include 1–2 keynotes, paper presentations (flexible length up to 12 pages in CEUR Workshop Proceedings format), a poster session, and an optional panel discussion. We will encourage interactive participation through discussion sessions and a prepared list of candidate topics for group discussion.

Target Audience. This workshop targets researchers and practitioners in natural language processing, information retrieval, human-computer interaction and cognitive science, especially those working on LLMs, dialogue systems, and user modelling. It welcomes participants from applied domains (e.g., healthcare, education, and enterprise search). We estimate an attendance of 25 to 40 participants, including paper authors, keynote speakers, and general attendees.

Workshop Schedule/Important Dates.

- Submission deadline: August 31, 2025
- Notification of acceptance: September 30, 2025
- Camera-ready version due: October 10, 2025

Time	Activity	Details
09:00 - 09:15	Opening & Welcome	Workshop introduction and overview
09:15 - 10:15	Keynote Talk	Invited Keynotes on Proactive Conversational IR
10:15 - 10:45	Coffee Break	Networking and discussions
10:45 - 12:00	Paper Presentations	Selected paper presentations
12:00 - 13:30	Lunch Break	Lunch and networking
13:30 - 15:00	Panel Discussion	Future directions and challenges
15:00 - 15:30	Coffee Break	Networking and discussions
15:30 - 17:00	Interactive Session	Hands-on activities and demos
17:00 - 17:15	Closing	Wrap-up and next steps

Table 1: Tentative Schedule for the ProActLLM workshop at CIKM 2025.

- Workshop date: November 14, 2025

5 Workshop Relevance

CIKM has long been a leading venue for work in information retrieval, conversational search, and applied NLP. With recent advances in LLMs, the community is well-positioned to explore proactive, context-aware conversational systems. ProActLLM aligns with CIKM 2025 themes — including information access, user interaction, evaluation, and generative AI — by offering a timely, focused forum on initiative-taking behaviors in LLM-based systems. It fills a gap in current offerings and will foster interdisciplinary exchange across IR, NLP, and HCI communities.

6 Past workshops by the organizers

The organisers of ProActLLM have been actively organising related workshops in the past, and here are the selected ones as examples:

(1) **UM-CIR**[1]: User Modelling in Conversational IR, hosted in SIGIR-AP 2024, (submission/acceptance/attendees: 5/3/15).
(2) **LLM-IGS**[2]: Large Language Models for Individuals, Groups and Society, hosted in WSDM 2024 and SIGIR 2024, (submission/acceptance/attendees: 7/6/25).
(3) **PASIR**[3]: Proactive and Agent-Supported Information Retrieval, hosted in CIKM 2022, (submitted/accepted/attendees: 4/3/5).
(4) **LLMIT**[4]: Large Language Models' Interpretation and Trustworthiness, hosted in CIKM 2023, (submissions/acceptance/attendees: 5/3/7).
(5) **LND4IR**[5]: Learning from Limited or Noisy Data for Information Retrieval in SIGIR 2018, (submission/acceptance/ attendees: 11/8/100+).

7 Related Workshops

While several related workshops have explored conversational AI and information retrieval (SCAI, NLP4ConvAI@ACL 2024, UM-CIR, MICROS@CIKM 2022, PASIR@CIKM2022, SUD@WSDM2021), our workshop is the first to focus specifically on the intersection of proactivity, conversational information seeking, and LLMs. By hosting it at CIKM, we engage the core IR and knowledge management community, while also attracting researchers from related areas. ProActLLM fills a unique gap not addressed by existing workshops, fostering a new cross-disciplinary community and advancing a timely, underexplored research direction.

8 Tentative Program Committee

(1) Jerome Ramos, University College London, UK
(2) Feng Xia, University of Sheffield, UK
(3) Hossein A. Rahmani, University College London, UK
(4) Xiao Fu, University College London, UK
(5) Hamish Clark, University of Glasgow, UK
(6) Suzan Verberne, Leiden University, the Netherlands
(7) Faegheh Hasibi, Radboud University, the Netherlands
(8) Chuan Meng, University of Amsterdam, the Netherlands
(9) Ebrahim Bagheri, University of Toronto, Canada
(10) Laura Dietz, University of New Hampshire, USA
(11) Mohammad Aliannejadi, University of Amsterdam, the Netherlands

9 Participation and Selection Process

Participation will be open to all registered attendees. We will solicit original research papers and perspective papers, reviewed through a double-blind process by the programme committee. Selection will be based on relevance, originality, technical quality, and clarity. Accepted papers will be presented as talks or posters.

10 Organisers

The organization team includes active IR and NLP researchers, with a common research interest in conventional information seeking, from both academia and industry.

- **Shubham Chatterjee (Lead Organiser)**: Assistant Professor, Missouri University of Science and Technology, USA. He focuses on information retrieval and retrieval augmented generation using large language models, which create systems that bridge complex data and human understanding.
- **Xi Wang (Lead Organiser)**: Lecturer (Assistant Professor) in Natural Language Processing at the University of Sheffield,

[1]UM-CIR: https://um-cir.github.io/
[2]LLM-IGS: https://llm-for-individuals-groups-and-society.github.io/2024_2
[3]PASIR: https://pasircikm2022.github.io/PASIRCIKM/
[4]LLMIT: https://gdebasis.github.io/llmit/
[5]LND4IR: https://lnd4ir.github.io/

UK. His research focuses on interactive natural language processing, including conversational AI, personalisation, RAG and related topics within interactive NLP.

- **Shuo Zhang**, Senior Research Scientist in Bloomberg AI, London. His interests include information retrieval, table retrieval, conversational AI, and text mining.
- **Sajad Ebrahimi**, M.A.Sc student in Computer Engineering at the University of Guelph. His research focuses on the generative ranking, diffusion models, and retrieval augmented generation.
- **Zhaochun Ren**, Associate Professor at Leiden University's LIACS, with primary interests in conversational AI, information retrieval, recommender systems, NLP, and large language models.
- **Debasis Ganguly**, Lecturer (Assistant Professor) in Data Science at the University of Glasgow. Former research staff at IBM Research Europe. His research covers data science, information retrieval, and explainability.
- **Gareth Jones**, Professor in the School of Computing at DCU. His research is centred on information retrieval, speech and multimedia search, personal archives, multilingual search, and search models.
- **Emine Yilmaz**, Professor of Computer Science at UCL, EPSRC and Alan Turing Institute fellow, Amazon Scholar, and co-founder of UCL spinout Humanloop. Her research spans information retrieval and NLP.
- **Hamed Zamani**, Associate Professor of CS and Associate Director of CIIR at UMass Amherst. His research focuses on neural IR, conversational search, and retrieval-enhanced machine learning.

References

[1] Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, and W Bruce Croft. 2019. Asking clarifying questions in open-domain information-seeking conversations. In *Proceedings of the 42nd international acm sigir conference on research and development in information retrieval*. 475–484.

[2] Yang Deng, Lizi Liao, Wenqiang Lei, Grace Hui Yang, Wai Lam, and Tat-Seng Chua. 2025. Proactive conversational ai: A comprehensive survey of advancements and opportunities. *ACM Transactions on Information Systems* 43, 3 (2025), 1–45.

[3] Zhengyuan Liu, Siti Salleh, Pavitra Krishnaswamy, and Nancy Chen. 2024. Context aggregation with topic-focused summarization for personalized medical dialogue generation. In *Proceedings of the 6th Clinical Natural Language Processing Workshop*. 310–321.

[4] Chuan Meng, Francesco Tonolini, Fengran Mo, Nikolaos Aletras, Emine Yilmaz, and Gabriella Kazai. 2025. Bridging the gap: From ad-hoc to proactive search in conversations. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 64–74.

[5] Chen Qu, Liu Yang, W Bruce Croft, Yongfeng Zhang, Johanne R Trippas, and Minghui Qiu. 2019. User intent prediction in information-seeking conversations. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*. 25–33.

[6] Chris Samarinas and Hamed Zamani. 2024. ProCIS: A benchmark for proactive retrieval in conversations. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 830–840.

[7] Xuan Zhang, Yang Deng, Zifeng Ren, See-Kiong Ng, and Tat-Seng Chua. 2024. Ask-before-Plan: Proactive Language Agents for Real-World Planning. In *EMNLP (Findings)*.

[8] Wayne Xin Zhao, Jing Liu, Ruiyang Ren, and Ji-Rong Wen. 2024. Dense text retrieval based on pretrained language models: A survey. *ACM Transactions on Information Systems* 42, 4 (2024), 1–60.