



Latest updates: <https://dl.acm.org/doi/10.1145/3726302.3730144>

DEMONSTRATION

Conversational Bibliographic Search

MARKUS NILLES, University of Trier, Trier, Rheinland-Pfalz, Germany

RALF SCHENKEL, University of Trier, Trier, Rheinland-Pfalz, Germany

Open Access Support provided by:

University of Trier



PDF Download
3726302.3730144.pdf
10 February 2026
Total Citations: 0
Total Downloads: 1668

Published: 13 July 2025

[Citation in BibTeX format](#)

SIGIR '25: The 48th International ACM SIGIR Conference on Research and Development in Information Retrieval
July 13 - 18, 2025
Padua, Italy

Conference Sponsors:
SIGIR

Conversational Bibliographic Search

Markus Nilles

Ralf Schenkel

nillesm@uni-trier.de

schenkel@uni-trier.de

Trier University

Trier, Germany

Abstract

Conversational Bibliographic Search is a conversational search engine designed to support users in retrieving scientific papers and authors within the domain of computer science. It enables natural language-based searches for data within the dblp computer science bibliography, which is enriched with data from Semantic Scholar. The system presents a novel user interface and bridges the gap between keyword-based search engines, faceted search systems, and generative conversational approaches. It enables users to access the latest publications, formulate intricate queries, as often required in scholarly research, and engage in multi-turn conversations to discover the most relevant results. Users can iteratively refine their queries and ask follow-up questions. Conversational Bibliographic Search actively supports the search process by posing clarification questions and providing suggestions.

CCS Concepts

- **Information systems → Search interfaces; Information retrieval query processing; Digital libraries and archives;** • **Human-centered computing → Natural language interfaces.**

Keywords

Conversational Information Retrieval, Conversational Search, Bibliographic Search, User Interface

ACM Reference Format:

Markus Nilles and Ralf Schenkel. 2025. Conversational Bibliographic Search. 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, 5 pages. <https://doi.org/10.1145/3726302.3730144>

1 Introduction

The rapid growth of scientific literature has made it increasingly difficult for researchers to efficiently find relevant papers and identify influential authors. Traditional search engines rely on one-shot keyword queries, often leading to incomplete or irrelevant results if users are unable to articulate their information needs. In addition, navigating and exploring large datasets using static search interfaces can be time-consuming, especially for users unfamiliar

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGIR '25, Padua, Italy

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1592-1/2025/07

<https://doi.org/10.1145/3726302.3730144>

with the latest terminology or trends, e.g., for less experienced undergraduate students. Faceted search systems for scientific publications offer a structured and interactive way to retrieve information. Researchers can refine search results by considering various dimensions, such as publication year, author, journal, topic, methodology, and citations. An alternative to traditional keyword-based search engines and faceted search systems are generative large language model (LLM) approaches for searching scientific publications. These approaches leverage advanced natural language processing to enhance the discovery and comprehension of research papers. Generative LLMs can interpret complex queries and allow multi-turn query refinement.

To leverage the advancements of these search engine types, we present *Conversational Bibliographic Search*, a conversational search engine designed to assist researchers in discovering scientific papers and authors through an intuitive, dialogue-based interface. Utilizing natural language understanding (NLU) and interactive question-answering techniques, our system enables users to engage in a conversational process to refine their queries, clarify ambiguous information, and dynamically explore related topics. While generative LLM approaches allow users to search in this manner, they face certain challenges that we address through our approach. One potential issue with generative LLM approaches is the risk of hallucination, where the model generates incorrect or non-existent references. Additionally, without transparency regarding how and on which data the model was trained, it may introduce biases. Furthermore, the computational cost of training and running such models can be high, especially considering the rapid publication of new papers daily.

Our system is showcased using computer science publications, based on data from dblp¹ and Semantic Scholar². It supports multiple search strategies such as natural language query formulation, query refinement in multiple turns, clarification questions, and providing query suggestions. Through conversations, our system makes it easier for scientists to find influential works and important authors in their research area. The user interface of the system enables users to find the most relevant publications and authors in the chat window. Additionally, users can access more search results in an additional side window if they wish to explore further.

In this paper, we describe the system architecture, the underlying technologies used, and the use cases that demonstrate its effectiveness in real-world research scenarios. We also highlight the interactive features that allow users to iteratively refine their search through natural conversation, making the system adaptable to different levels of search expertise and research goals.

¹<https://dblp.org>

²<https://www.semanticscholar.org>

A demonstration video of Conversational Bibliographic Search is available at <https://video.uni-trier.de/Panopto/Pages/Viewer.aspx?id=479a48e6-2481-4d96-8e65-b28801128cee>.

2 Related Work

The evolution of natural language processing and the demand for more intuitive search interactions have paved the way for the development of conversational search systems. Unlike traditional keyword-based search engines, these systems enable multi-turn dialogues, allowing users to refine their queries and resolve ambiguities iteratively. Prominent examples include ChatGPT with its search option³, Microsoft's Copilot⁴, or Perplexity.ai⁵ which use the power of context handling and dialogue management to create a more interactive and user-friendly search experience.

We chose to design a custom architecture for our conversational search system due to the specific challenges in the field of bibliographic search. However, there's an open-source framework called Macaw [10] that could be used to build a system similar to Conversational Bibliographic Search. Macaw supports multi-turn, multi-modal, and mixed-initiative interactions across various tasks, including document retrieval, question answering, and recommendation. While there exist further works that deal with individual components of a conversational information retrieval system, e.g. user intent classification and slot filling (Louvan and Magnini [8]) or response generation (Lajewska and Balog [6]), there is no system that focuses explicitly on the implementation of a conversational information retrieval system for bibliographic metadata that handle the unique challenges of scholarly search. In contrast to the data of the previous mentioned conversational information retrieval systems, the data of bibliographic metadata is extensive but domain-specific and search queries have a vocabulary corresponding to the domain.

Current search engines for bibliographic metadata, such as ResearchGate⁶, Semantic Scholar or Google Scholar allow only keyword-based searches. Kreutz et al. [5] presented SchenQL, a query language for bibliographic metadata that allows users to make their queries more easily and precisely. OpenScholar [1], PaperQA [7], and PaperQA2 [9] utilize retrieval-augmented language models to improve access to scientific literature.

3 Architecture

The system's architecture, as illustrated in Figure 1, consists of five main components. It follows a modular design similar to those architectures employed in the systems of Zeng et al. [11] and Kemper et al.[4] for natural language understanding, dialogue state updating, action selection, and response generation. The User Intent Classification & Slot Filling Module updates the current conversational state by considering the previous conversational state and extracting the user intent and slot value pairs from the current user query. The Search Module queries a Lucene index of the bibliographic data to meet the user's current information requirement. The Conversational Module uses the current conversational state

³<https://chatgpt.com/search>

⁴<https://copilot.microsoft.com/>

⁵<https://www.perplexity.ai/>

⁶<https://www.researchgate.net>

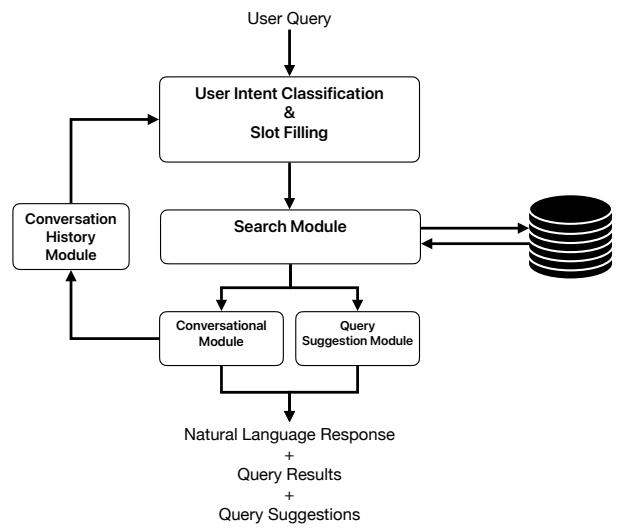


Figure 1: The architecture of Conversational Bibliographic Search with its five components: User Intent Classification & Slot Filling, Search Module, Conversational Module, Query Suggestion Module, Conversation History Module

and the retrieved results to generate a natural language response. The Query Suggestion Module uses the current conversational state and the retrieved results to provide suggestions to users for their subsequent queries. The Conversational History Module tracks the conversational state, ensuring it's considered in the next turn of the conversation.

3.1 User Intent Classification & Slot Filling

The User Intent Classification & Slot Filling Module has two tasks: First, it identifies the user's intent based on their utterances; second, it extracts specific slot values from those utterances. Based on the data from dblp and Semantic Scholar, we defined a set of user intents and a set of slots. User intents can be publications, authors, or specific details about a publication or an author. We also added user intents for queries where the user wants to refine their previous queries. Additionally, we included intents that allow users to specify their preferred order of retrieved publications or authors. Users can ask for the most relevant, most cited, newest, or oldest publications. The slots define the user-input criteria that must be met to satisfy their information requirement. For instance, these slots could include a keyword, a title, a person's name, or the venue of a publication. Additionally, users can filter the results by specifying a particular year, a range of years when the publication was published, or the minimum number of citations it should have. We formulated user utterances for each of these user intents and slots and utilized ChatGPT to generate additional user utterances for the training of a joint intent classification and slot filling model based on BERT [2, 3]. Currently the model is trained with 2937 user utterances.

The previous conversational state is updated based on the determined user intent and slot values to enable refinements over

multiple turns. For instance, a user could search for publications on a specific topic in the first query and request a summary of a retrieved publication in the second query without explicitly mentioning it. Another option for the user to alter the current conversational state is to select a publication or a person to focus on. Each publication or person visible in the user interface has a target icon button that can be clicked. When the target button is clicked, the publication or person is added to the conversational state, allowing the user to directly formulate a query about that object without providing any slot values in their query. The target icon also indicates the publication or person being focused on. If the current object being focused on isn't the desired one, the user can select another object or formulate a new natural language query, and the system automatically determines whether the new utterance refers to the object or if the user wants to switch the context.

3.2 Search Module

The Search Module queries a Lucene index containing the bibliographic data. The conversational state updated by the User Intent Classification & Slot Filling determines how the query will be build. The Search Module checks for specific slots, such as the name of a person, journals, and venues, if the system requires clarification. For instance, if a user searches for publications by an author's name and there are multiple individuals with the same name, a clarification question is presented to the user. The system provides options for the user to select the desired individual.

A query for a conversational state consists of multiple subqueries. For each slot value pair in the conversational state, a subquery is added to the query. Each slot's subquery will then search in one or more fields of the index.

After the retrieval, the results are presented to the user. The user has the option to view all the details of the publications and individuals by opening a dedicated detail page. To access this page, they can click on the arrow icon associated with the object. The user can navigate to the pages of related publications or individuals by clicking on their titles or names. Additionally, the user can always return to the chat window. If they're interested in continuing the conversation by asking questions about an object they've discovered, they can simply click the target icon next to the object.

3.3 Conversational Module

The Conversational Module formulates the natural language answer of the system using predefined templates. The determined user intent and slots and the retrieved results are inserted in the templates. Since it is not useful to present all the retrieved results in the chat, only the top results are presented.

If the user request is too vague, the Conversational Module assists the user in refining the query for greater precision. In cases of ambiguous queries, the system prompts the user with a clarification question, allowing them to select from the suggested options. For instance, when there are multiple authors with the same name, a clarification question is asked. Additionally, when the system is uncertain about which journal or venue they want to see publications from, a clarification question is prompted.

During the conversation, the system offers explanations and assists the user in interacting with it and understanding its supported

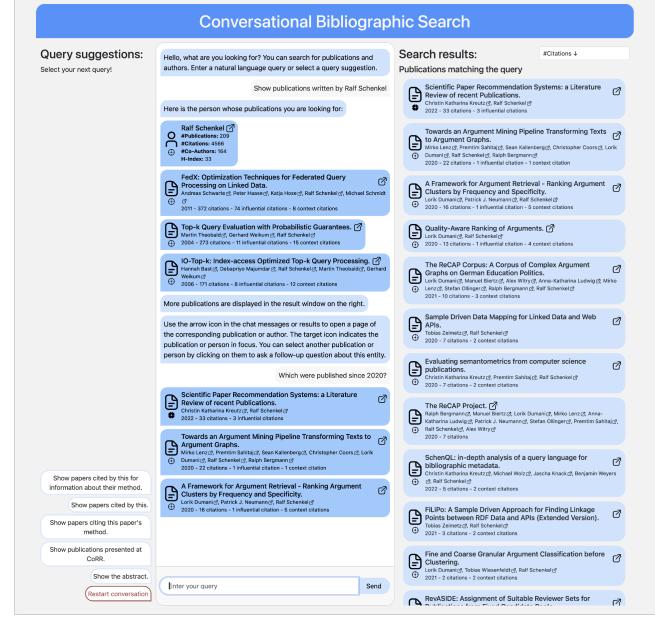


Figure 2: Main User Interface of Conversational Bibliographic Search.

features. These explanations are provided only when necessary, such as when the user first sees results.

3.4 Query Suggestion Module

The Query Suggestion Module provides users with suggestions for their next query. These suggestions are generated based on the current conversational state and the results currently displayed. Similar to natural language responses, query suggestions are created using templates. The set of suggestions for the next turn includes queries for details about publications or persons, queries for publications related to similar topics, publications presented at the same venue, or publications published in the same journal.

3.5 Conversation History Module

The Conversation History Module stores the user queries, the system responses, and the query results to update the dialogue state after each turn. The dialogue state is updated by the User Intent Classification & Slot Filling Module. The user utterances can also be used to analyze the user's search tasks and to improve the system.

4 Demonstration

Conversational Bibliographic Search is a web-based application that presents a unique user interface comprising three distinct pages: a main interface for conducting searches, a detailed page for publications, and a detailed page for authors.

The main interface, shown in Figure 2, is divided in three windows displayed in a side-by-side configuration. The center window serves as a chat interface, facilitating a natural language dialogue between the system and the user. Users can type their queries into a dedicated text field and the system will respond accordingly. An alternative to typing the query in the text field is provided in the

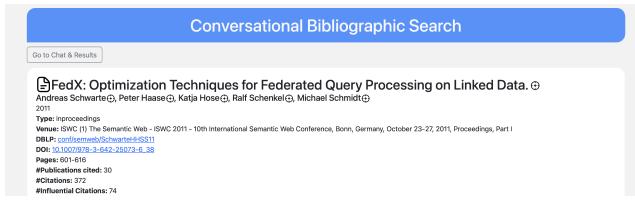


Figure 3: Detail page of a publication.

left-hand window. This window offers query suggestions, which the user can select by clicking on them. These suggestions are generated based on the current conversational state and the results displayed. The suggested queries may request additional publications from the same journals or venues as the results, or they may ask for publications cited by the displayed publications or co-authors of the individuals mentioned in the results. The right window functions as a results display, showing additional information such as relevant results, publications, or authors. In cases where multiple results are retrieved, presenting all of them in the results window is more effective than in the chat interface because it allows users to scan the information more efficiently. Additionally users are able to select the order in which the results are presented.

In the example conversation depicted in Figure 2, the system initiated the conversation and the user asked for publications of the author *Ralf Schenkel*. The system analyzed the query and retrieved his most cited publications. In the chat window, the system responds in natural language and presents a message that includes information about the author and messages with details of the top three publications. Since it was the first time results were retrieved, the system provides explanations of potential interactions with the person and publication objects. Clicking the arrow icon opens a page containing detailed information about the corresponding publication or person. The target icon allows a publication or person to be set in focus, enabling the possibility a follow-up query about it.

The user refined their query by searching only publications published since 2020. The system updated the results, presenting the user with the most cited publications authored by the same author since that year. Notably, the author has written more than just the three publications displayed in the chat window since 2020. Therefore, the user has the opportunity to explore all of the author's publications in the results window. The user has several options to further engage with these publications: they can open the detailed page of any of these publications or their authors, or they can shift the focus to any of these publications or authors to ask a question about them. The system highlights the object currently in focus by displaying a filled target icon. Users can request, for instance, a summary or the abstract, or information about the publications cited by this publication or the publications that cite this publication. The user could also choose to click on a query suggestion displayed in the left window.

If a user navigates to the detailed page of a publication (Figure 3) or a person, they can access all the information stored in the database. On the side of the publication's titles or author's names, there's a target button that allows the user to return to the main interface to continue the conversation with the publication or person in focus.

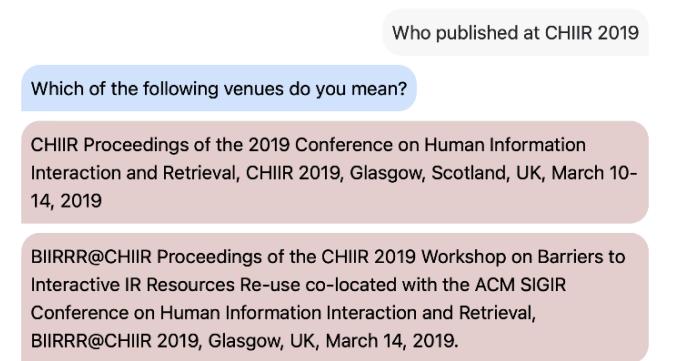


Figure 4: Clarification question asked by the system.

The system is designed to recognize instances requiring clarification and can proactively prompt the user with a clarification question. Figure 4 shows an example of this, where the user requests publications from a specific conference, yet ambiguity arises due to multiple conferences containing "CHIIR 2019". The system responds with a clarification question, prompting the user to select the desired conference, on which the relevant publications are displayed. Clarification questions are also prompted when a user requests publications by an individual with a common name, resulting in multiple potential matches. In this scenario, the system presents the user with a list of options, allowing them to identify and select the correct individual to retrieve the desired publications.

5 Future Work

Conversational Bibliographic Search is a conversational information retrieval system designed to search for scientific papers and authors in the field of computer science. The current version allows users to formulate queries based on a predefined set of intents and slot values, refine those queries through multiple interactions, and resolve ambiguities through answering clarification questions. It also suggests queries for the next interaction to help the user refine their query or search related publications or authors. The user interface presented combines a conversational chat interface with an extended result list for the retrieved publications and authors. Users can view the most relevant results directly in the chat, while still having the option to explore more results in the results window located on the side of the chat window.

In the future, we plan to enhance the system with additional features. We aim to expand the current range of user intents and slot options to include a broader set of possibilities. Through user studies, we aim to identify additional features that users would find valuable in a conversational literature retrieval system. We will also evaluate the effectiveness and user experience of the system by comparing it to other search engines, including keyword-based systems, faceted search interfaces, systems using generative LLM approaches and systems that use alternative query languages such as SQL. In addition, we will investigate which user groups benefit the most from a conversational search interface compared to other search methods.

References

- [1] Akari Asai, Jacqueline He, Rulin Shao, Weijia Shi, Amanpreet Singh, Joseph Chee Chang, Kyle Lo, Luca Soldaini, Sergey Feldman, Mike D'Arcy, David Wadden, Matt Latzke, Minyang Tian, Pan Ji, Shengyan Liu, Hao Tong, Bohao Wu, Yanyu Xiong, Luke Zettlemoyer, Graham Neubig, Daniel S. Weld, Doug Downey, Wen-tau Yih, Pang Wei Koh, and Hannaneh Hajishirzi. 2024. OpenScholar: Synthesizing Scientific Literature with Retrieval-augmented LMs. *CoRR* abs/2411.14199 (2024). <https://doi.org/10.48550/ARXIV.2411.14199> arXiv:2411.14199
- [2] Qian Chen, Zhuo Zhuo, and Wen Wang. 2019. BERT for Joint Intent Classification and Slot Filling. *CoRR* abs/1902.10909 (2019). arXiv:1902.10909 <http://arxiv.org/abs/1902.10909>
- [3] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, Jill Burstein, Christy Doran, and Thamar Solorio (Eds.). Association for Computational Linguistics, 4171–4186. <https://doi.org/10.18653/V1/N19-1423>
- [4] 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, SIGIR 2024, Washington DC, USA, July 14–18, 2024*, Grace Hui Yang, Hongning Wang, San Han, Claudia Hauff, Guido Zuccon, and Yi Zhang (Eds.). ACM, 2786–2790. <https://doi.org/10.1145/3626772.3657670>
- [5] Christian Katharina Kreutz, Michael Wolz, Jascha Knack, Benjamin Weyers, and Ralf Schenkel. 2022. SchenQL: in-depth analysis of a query language for bibliographic metadata. *Int. J. Digit. Libr.* 23, 2 (2022), 113–132. <https://doi.org/10.1007/s00799-021-00317-8>
- [6] Weronika Lajewska and Krisztian Balog. 2023. Towards Filling the Gap in Conversational Search: From Passage Retrieval to Conversational Response Generation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, CIKM 2023, Birmingham, United Kingdom, October 21–25, 2023*, Ingo Frommholz, Frank Hopfgartner, Mark Lee, Michael Oakes, Mounia Lalmas, Min Zhang, and Rodrygo L. T. Santos (Eds.). ACM, 5326–5330. <https://doi.org/10.1145/3583780.3615132>
- [7] Jakub Lála, Odhran O'Donoghue, Aleksandar Shtedritski, Sam Cox, Samuel G. Rodrigues, and Andrew D. White. 2023. PaperQA: Retrieval-Augmented Generative Agent for Scientific Research. *CoRR* abs/2312.07559 (2023). <https://doi.org/10.48550/ARXIV.2312.07559> arXiv:2312.07559
- [8] Samuel Louvan and Bernardo Magnini. 2020. Recent Neural Methods on Slot Filling and Intent Classification for Task-Oriented Dialogue Systems: A Survey. In *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8–13, 2020*, Donia Scott, Núria Bel, and Chengqing Zong (Eds.). International Committee on Computational Linguistics, 480–496. <https://doi.org/10.18653/V1/2020.COLING-MAIN.42>
- [9] Michael D. Skarbinski, Sam Cox, Jon M. Laurent, James D. Braza, Michaela M. Hinks, Michael J. Hammerling, Manvitha Ponnappati, Samuel G. Rodrigues, and Andrew D. White. 2024. Language agents achieve superhuman synthesis of scientific knowledge. *CoRR* abs/2409.13740 (2024). <https://doi.org/10.48550/ARXIV.2409.13740> arXiv:2409.13740
- [10] Hamed Zamani and Nick Craswell. 2020. Macaw: An Extensible Conversational Information Seeking Platform. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25–30, 2020*, Jimmy X. Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu (Eds.). ACM, 2193–2196. <https://doi.org/10.1145/3397271.3401415>
- [11] Yankai Zeng, Abhiram Rajasekharan, Parth Padalkar, Kinjal Basu, Joaquin Arias, and Gopal Gupta. 2024. Automated Interactive Domain-Specific Conversational Agents that Understand Human Dialogs. In *Practical Aspects of Declarative Languages - 26th International Symposium, PADL 2024, London, UK, January 15–16, 2024, Proceedings (Lecture Notes in Computer Science, Vol. 14512)*, Martin Gebser and Ilya Sergey (Eds.). Springer, 204–222. https://doi.org/10.1007/978-3-031-52038-9_13