Conversational Information Seeking: Theory and Evaluation

CHIIR 2022 Half Day Tutorial

Mohammad Aliannejadi
University of Amsterdam
Amsterdam, The Netherlands
m.aliannejadi@uva.nl

Johanne R. Trippas
University of Melbourne
Melbourne, Australia
johanne.trippas@unimelb.edu.au

ABSTRACT

Research in conversational information seeking (CIS) is moving very rapidly in various directions such as user interaction, system design, and evaluation. The tutorial focuses on the theoretical foundations and information-seeking processes for CIS, as well as their evaluation. The tutorial aims to introduce and communicate CIS research to the community and discuss it from different perspectives, such as theoretical modelling, evaluation, and user simulation. Also, it aims at gathering researchers and practitioners interested in this research direction for discussions, idea communications, and research promotions.

ACM Reference Format:
Mohammad Aliannejadi and Johanne R. Trippas. 2022. Conversational Information Seeking: Theory and Evaluation: CHIIR 2022 Half Day Tutorial. In *Proceedings of the 2022 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '22), March 14–18, 2022, Regensburg, Germany.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3498366.3505843

1 INTRODUCTION AND MOTIVATION

Conversational information seeking (CIS), the process in which people search for information through conversations, is rising in popularity. Furthermore, CIS has been suggested as one of the next paradigm shifts in how people search for information [5, 8]. Despite several advances in the field, not much effort has been made to present different aspects of CIS. In this tutorial, we focus on two essential research aspects of CIS: theory and evaluation.

2 SYLLABUS AND STRUCTURE

We aim to present a half-day tutorial (3 hours). We will incorporate interactive activities by using software such as *Poll Everywhere*[1] to overcome online interaction barriers.

The tutorial will be split into two main sessions with different core topics delivered by both presenters.

[1]https://www.polleverywhere.com/

CHIIR '22, March 14–18, 2022, Regensburg, Germany

ACM ISBN 978-1-4503-9186-3/22/03…$15.00
https://doi.org/10.1145/3498366.3505843

2.1 Part I: Introduction to CIS

This first part of the tutorial provides an introduction and overview to searching through conversational interactions and is divided into three main subsections.

Subsection I: Theoretical concepts in CIS. We introduce what it means to have a conversational system, including topics such as natural language and conversational interactions. Then, we introduce the concept of CIS and present its different definitions. We will introduce these existing definitions based on prior human conversations, theoretical instigation, and empirical studies. We will introduce how researchers have attempted to clarify what conversational search is from a user and system point of view [6, 13], a cognitive perspective [16], or through user–system interactivity [19]. We explain how CIS systems will need to shift beyond the "command and control" paradigm. Lastly, we elaborate on the different interaction modalities (i.e., text-only, audio-only, or multi-modal interactions).

Subsection II: Existing conversational systems versus CIS. We present the background to relevant conversational systems types and introduce a historical overview of conversational systems, including spoken dialogue systems, voice user interfaces, chatbots, and live chat support. We illustrate the differences between these existing groups of conversational systems in relationship to CIS.

Subsection III: Aligning existing information-seeking processes with CIS. We present the possible interactions and functional goals for CIS through the information search behaviours lens [11]. First, we examine traditional search interactions (query formulation, result presentation and answer organization, and query reformulation or refinements). We will pay extra attention to result presentation as a critical element of transferring knowledge in a highly interactive setting such as CIS. Last, we investigate fundamental non-search interactions such as discourse management and grounding and how they play a crucial role in CIS.

2.2 Part II: CIS Evaluation

The second part of the tutorial focuses on various challenges evaluating researchers face for CIS. This part of the tutorial is divided into three subsections.

Subsection I: Conversational Passage Ranking. We present the evaluation framework presented at the TREC Conversational Assistant Track (CAsT) [9], and how the organizers formulate conversational passage ranking as a single-initiative conversational system.

Subsection II: Mixed-Initiative Conversations. We present the approaches presented to evaluate mixed-initiative conversations in

an offline setting using automatic [4] and human evaluation metrics [3, 10], followed by the details of the ConvAI3 shared task [2], and how human-in-the-loop techniques can be used to facilitate the evaluation of generative models.

Subsection III: User Simulation. We present a line of research on user simulation for CIS evaluation. User simulation has been used to evaluate recommender systems [21], where the nature of the available data is structured. We present the CoSearcher [14] that aims to simulate users in a mixed-initiative setup and facilitate evaluation. Furthermore, we discuss the limitations of CoSearcher, such as its limitation on generating natural language user answers and present USi [15] that addresses such limitations, aiming to provide a more practical user simulation. We finally present an analytical work [1] that utilizes user simulations to study various conversational strategies and mixed initiatives and their trade-offs.

3 TARGET AUDIENCE

This tutorial is for anyone interested in better understanding CIS. We aim to keep the introduction to CIS broad and is indented as a starting point to research in CIS, primarily aimed at graduate students.

3.1 Learning Outcomes

The tutorial aims to educate the audience about the basic concepts of CIS and the techniques used so far to evaluate such complex systems. Our goal is to provide a broad view of various approaches, and challenges researchers face while developing and evaluating CIS.

4 PRESENTERS

Mohammad Aliannejadi (http://aliannejadi.com) Dr Aliannejadi is a post-doctoral researcher at the *IRLab, University of Amsterdam*, The Netherlands. He obtained his PhD in Informatics from Universitá della Svizzera italiana (USI), Switzerland, under the supervision of Professor Fabio Crestani. During his PhD, he visited the CIIR group at UMass Amherst, where he worked with Professor W. Bruce Croft. Recently his work has focused on user simulation for mixed-initiative conversational systems and the generation of clarifying questions. He has been a *SIGIR* Student Liaison and has co-organised workshops on mixed-initiative conversations [12] and search-based conversational systems. He has given several lectures on conversational search and mixed initiatives at the University of Amsterdam and SIKS[2].

Johanne Trippas (http://johannetrippas.com) Dr Trippas is a *Doreen Thomas Research Fellow* at the *University of Melbourne*, Australia. The position is a competitive independent fellowship to pursue the research of her choice. She obtained her PhD in Computer Science from *RMIT University*, Australia, under the supervision of Professor Mark Sanderson, Professor Lawrence Cavedon, and Dr Damiano Spina. She was awarded the RMIT University Deputy Vice-Chancellor's Higher Degree by Research Prize for her doctoral work and thesis focusing on spoken conversational search [16]. Recently, her work has focused on developing next-generation capabilities for intelligent systems, including spoken conversational search, digital assistants in a cockpit, and Artificial Intelligence to identify cardiac arrests. She has been a *SIGIR* Student Liaison, has co-organised several CHIIR tutorials and workshops [7, 17, 18, 20], and the Russian Summer School on Information Retrieval 2020 (cancelled due to COVID-19). Furthermore, Johanne is actively involved in the broader SIGIR community and was recently appointed as vice-chair of the *SIGIR* Artifact Evaluation Committee.

[2]http://siks.nl

REFERENCES

[1] Mohammad Aliannejadi, Leif Azzopardi, Hamed Zamani, Evangelos Kanoulas, Paul Thomas, and Nick Craswel. 2021. Analysing Mixed Initiatives and Search Strategies during Conversational Search. In *CIKM*.

[2] Mohammad Aliannejadi, Julia Kiseleva, Aleksandr Chuklin, Jeff Dalton, and Mikhail S. Burtsev. 2020. ConvAI3: Generating Clarifying Questions for Open-Domain Dialogue Systems (ClariQ). *CoRR* abs/2009.11352 (2020).

[3] Mohammad Aliannejadi, Julia Kiseleva, Aleksandr Chuklin, Jeffrey Dalton, and Mikhail S. Burtsev. 2021. Building and Evaluating Open-Domain Dialogue Corpora with Clarifying Questions. In *EMNLP*.

[4] Mohammad Aliannejadi, Hamed Zamani, Fabio Crestani, and W. Bruce Croft. 2019. Asking Clarifying Questions in Open-Domain Information-Seeking Conversations. In *SIGIR*. ACM, 475–484.

[5] Avishek Anand, Lawrence Cavedon, Hideo Joho, Mark Sanderson, and Benno Stein. 2020. Conversational Search (Dagstuhl Seminar 19461). *Dagstuhl Reports* 9, 11 (2020), 34–83.

[6] Leif Azzopardi, Mateusz Dubiel, Martin Halvey, and Jeffrey Dalton. 2018. Conceptualizing agent-human interactions during the conversational search process. In *SIGIR 2nd International Workshop on Conversational Approaches to Information Retrieval (CAIR'18)*. 8 pages.

[7] George Buchanan, Dana McKay, Charles LA Clarke, Leif Azzopardi, and Johanne R. Trippas. 2020. Made to Measure: A Workshop on Human-centred metrics for information seeking. In *Proceedings of Conference on Information Interaction and Retrieval (CHIIR)*. 484–487.

[8] J. Shane Culpepper, Fernando Diaz, and Mark D. Smucker. 2018. Research Frontiers in Information Retrieval: Report from the Third Strategic Workshop on Information Retrieval in Lorne (SWIRL 2018). *SIGIR Forum* 52, 1 (2018), 34–90.

[9] Jeffrey Dalton, Chenyan Xiong, and Jamie Callan. 2020. CAsT 2020: The Conversational Assistance Track Overview. In *TREC (NIST Special Publication, Vol. 1266)*. National Institute of Standards and Technology (NIST).

[10] Margaret Li, Jason Weston, and Stephen Roller. 2019. ACUTE-EVAL: Improved Dialogue Evaluation with Optimized Questions and Multi-turn Comparisons. *CoRR* abs/1909.03087 (2019).

[11] Garry Marchionini. 1997. *Information Seeking in Electronic Environments*. Cambridge University Press.

[12] Ida Mele, Cristina Ioana Muntean, Mohammad Aliannejadi, and Nikos Voskarides. 2021. MICROS: Mixed-Initiative ConveRsatiOnal Systems Workshop. In *ECIR (2) (Lecture Notes in Computer Science, Vol. 12657)*. Springer, 710–713.

[13] Filip Radlinski and Nick Craswell. 2017. A theoretical framework for conversational search. In *Proceedings of Conference on Information Interaction and Retrieval (CHIIR)*. 117–126.

[14] Alexandre Salle, Shervin Malmasi, Oleg Rokhlenko, and Eugene Agichtein. 2021. Studying the Effectiveness of Conversational Search Refinement Through User Simulation. In *ECIR (1) (Lecture Notes in Computer Science, Vol. 12656)*. Springer, 587–602.

[15] Ivan Sekulic, Mohammad Aliannejadi, and Fabio Crestani. 2022. Evaluating Mixed-initiative Conversational Search Systems via User Simulation. In *WSDM*.

[16] Johanne R. Trippas. 2019. *Spoken Conversational Search: Audio-only Interactive Information Retrieval*. Ph. D. Dissertation. RMIT, Melbourne.

[17] Johanne R. Trippas and David Maxwell. 2021. The PhD Journey: Reaching Out and Lending a Hand. In *Proceedings of Conference on Information Interaction and Retrieval (CHIIR)*. 492–494.

[18] Johanne R. Trippas and David Maxwell. 2022. First Early Career Researchers' Roundtable for Information Access Research. In *Proceedings of Conference on Information Interaction and Retrieval (CHIIR)*.

[19] Johanne R. Trippas, Damiano Spina, Paul Thomas, Mark Sanderson, Hideo Joho, and Lawrence Cavedon. 2020. Towards a Model for Spoken Conversational Search. *Information Processing & Management* 57, 2 (2020), 102162.

[20] Johanne R. Trippas, Paul Thomas, Damiano Spina, and Hideo Joho. 2020. Third CAIR Workshop. In *Proceedings of Conference on Information Interaction and Retrieval (CHIIR)*. 492–494.

[21] Shuo Zhang and Krisztian Balog. 2020. Evaluating Conversational Recommender Systems via User Simulation. In *KDD*. ACM, 1512–1520.