2nd Workshop on Information Retrieval for Understudied Users (IR4U2) - Bridging User-centered AI with IR: Making Information Retrieval Accessible for All

Mauro, Noemi; Cossatin, Angelo Geninatti; Pera, Maria Soledad; Cena, Federica; Landoni, Monica; Huibers, Theo; Murgia, Emiliana

DOI
10.1145/3726302.3730356

Publication date
2025

Document Version
Final published version

Published in
SIGIR '25: Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval

2nd Workshop on Information Retrieval for Understudied Users (IR4U2) - *Bridging User-centered AI with IR: Making Information Retrieval Accessible for All*

Noemi Mauro
noemi.mauro@unito.it
University of Turin
Turin, Italy

Angelo Geninatti Cossatin
angelo.geninatticossatin@unito.it
University of Turin
Turin, Italy

Maria Soledad Pera
m.s.pera@tudelft.nl
TU Delft
Delft, Netherlands

Federica Cena
federica.cena@unito.it
University of Turin
Turin, Italy

Monica Landoni
monica.landoni@usi.ch
Università della Svizzera Italiana
Lugano, Switzerland

Theo Huibers
t.w.c.huibers@utwente.nl
University of Twente
Enschede, The Netherlands

Emiliana Murgia
emiliana.murgia@unige.it
Università di Genova
Genova, Italy

Abstract

The **Workshop on Information Retrieval for Understudied Users** (IR4U2) serves as a platform to highlight information retrieval (IR) research that directly impacts often *understudied* user groups. The second IR4U2 workshop focuses on a user-centred AI perspective, which is vital for informing the design, development, and assessment of information retrieval systems that thoughtfully address the diverse needs of understudied populations, ensuring genuine accessibility and inclusivity. The objectives of IR4U2 are: (1) to build community and awareness by sharing AI and IR developments that serve underrepresented user groups in this research area; (2) to identify challenges and open issues along with lessons learned and challenges inherent to this area of research; and (3) to spark discussions that establish common frameworks for future research.

CCS Concepts

• **Information systems** → **Recommender systems**; **Users and interactive retrieval**; • **Human-centered computing**;

Keywords

Understudied users, Information Access, Information Retrieval

ACM Reference Format:

Noemi Mauro, Angelo Geninatti Cossatin, Maria Soledad Pera, Federica Cena, Monica Landoni, Theo Huibers, and Emiliana Murgia. 2025. 2nd Workshop on Information Retrieval for Understudied Users (IR4U2) - *Bridging User-centered AI with IR: Making Information Retrieval Accessible for All*. 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/10.1145/3726302.3730356

SIGIR '25, July 13–18, 2025, Padua, Italy

ACM ISBN 979-8-4007-1592-1/2025/07
https://doi.org/10.1145/3726302.3730356

1 Introduction and Motivation

Information retrieval (**IR**) is the dominant form of information access [5]. At the same time, Artificial Intelligence (**AI**) is reshaping the IR landscape, with all its power as well its weaknesses. Over the past few decades, developments in search, recommender, and question-answering systems, along with complementary IR tasks such as clustering, filtering, and text processing, have been well studied and documented; so are the efforts that help IR keep up with the fast-paced era of Large Language Models, AI, new modalities of interactions, and intelligent platforms enabling information access. We see, however, that works at SIGIR, ECIR, RecSys, and other related venues for the most part aid or study the so-called "classical" users, e.g., English-speaking adults with good literacy skills by keeping perspectives and biases inherent in the mainstream dominating Western culture.

Nowadays, there are more than 5 billion Internet users worldwide[1], seeking access or being exposed to information about travel, news, education, and research on a daily basis[2]. As the core area driving innovation, the IR community must ask itself whether current (IR) technologies genuinely support (the right to information for) all users. For instance, do popular search engines consider and cater for individuals with low literacy levels or biases affecting search results? Are recommendation algorithms on platforms like TikTok or YouTube inadvertently exposing children to harmful content? Are the AI technologies that drive IR solutions inclusive and fairly representative of diverse user populations? Are there any techniques to make AI in IR more accessible to users, regardless of their context? Indeed, some IR research exists focusing on improving access for diverse groups–such as older adults, children,

[1] https://www.statista.com/statistics/617136/digital-population-worldwide/
[2] https://www.statista.com/statistics/1387376/internet-using-global-reasons-by-age/

and those with disabilities. Still, these efforts are sporadic [1–4, 6–9, 11, 14] and often sidelined as addition to mains stream research. The IR community has started addressing fairness and bias, yet more work is needed to understand how IR technology can better serve all users.

Enhancing information access requires considering the environments where these technologies are used and the unique needs of users. The community is making strides through domain-specific IR initiatives in domains like medical [13] and legal [12]. However, the diverse needs of users are often overlooked. We argue that to effectively serve a varied population, we must adopt a human-centered approach that emphasizes real user needs and fosters cross-disciplinary dialogue [10].

The 2^{nd} Workshop on Information Retrieval for Understudied Users (IR4U2) aims to unite researchers and practitioners committed to advancing knowledge in this area, enhancing visibility for understudied groups within the IR community by co-locating with SIGIR. By sharing challenges and concerns researchers could ponder on the best approaches to adopt by possibly taking advantage of the role AI is playing in facilitating research as for instance in fast prototyping or user simulation. Moreover, this year's edition aims to focus on bridging User-centered AI with IR to foster contributions that put the different types of users, especially the understudied ones, at the center of the design of both AI-IR algorithms and their user interfaces.

While the first edition successfully established foundational discussions and community connections, the work of making IR systems truly inclusive remains ongoing. A second edition at SIGIR would build on this momentum, addressing identified challenges around user-centered AI, expanding the community of researchers and practitioners working in this space, and developing concrete frameworks for future research. Moreover, SIGIR's position as a premier IR venue offers an ideal venue to increase the visibility and impact of research focused on understudied users, helping bridge the gap between theoretical advances and practical applications that benefit diverse user groups. More so, this workshop is related to the *low-resource environments* track at SIGIR 2025[3]: as indeed being part of an understudied community often limits available resources for research effort, overall involvement and the development of specific tools. By addressing these gaps, we can make (IR) technology *more accessible to everyone.*

2 Theme and Purpose

This year, we seek to raise awareness to the topic of "*User-Centered AI in IR for understudied users*". User-centered AI in IR systems must be thoughtfully engineered to address the diverse needs of understudied populations, ensuring genuine accessibility and inclusivity. For individuals with disabilities, AI-powered systems can deliver customized adaptations, including speech-to-text conversion, detailed image descriptions, and dynamic sign language interpretation.

To achieve effective user-centered design, developers must conduct thorough user research through interviews, surveys, and direct observation of varied user demographics. Consider visually impaired users, who require specific IR accommodations such as optimized text dimensions, appropriate button sizing, and carefully calibrated color contrast in interface layouts. From an algorithmic perspective, IR systems should incorporate user interaction patterns specific to different user groups. This means developing adaptive ranking algorithms that for example adjust result relevance based on group-specific interaction patterns, account for varying levels of digital literacy and accessibility needs and implement personalized weighting schemes for different user segments. This dual approach–combining interface adaptability with algorithmic customization–ensures that IR systems serve all users effectively, regardless of their abilities or backgrounds.

The goals for this 2^{nd} IR4U2 are to:

- Build, extend and maintain a community involving multidisciplinary Ph.D. students, academics, and industry practitioners aiming to advance understanding of core (AI for) IR technology and its impact on a broad range of understudied users.
- Take inventory of the different kinds of use cases, research work, ongoing efforts, and existing resources while reflecting on the progress made since the last IR4U2 iteration.
- Define guidelines to promote user-centric design of IR systems.
- Outline a general approach to facilitate research endeavors driven by and at the service of specific user groups.

2.1 Workshop Plan

We will elicit contributions that can help outline a snapshot of the current works in this area and prompt directions for the future. Their topics center on, though are not restricted to: User-centered design of IR algorithms and interface for understudied; The use of AI for enhancing IR systems for understudied, focusing on positive and negative impact; User modeling to enable IR and recommendation technologies tailored to understudied populations; Data collection and benchmark development of IR catering to understudied populations; IR applications targeting understudied populations; IR-related technology (clustering, classification, text processing, text complexity) and their impact on understudied populations; Different perspectives of evaluation and ethical issues associated with IR and AI technologies for understudied users; e.g., their right to be represented, acknowledged, have access, and be served by existing and to-be-developed systems.

In this iteration of the workshop, we consider two types of submissions:

Peer-review. Contributions fitting this category will be peer-reviewed and include **(a) research papers** presenting empirical explorations, user studies, and/or new algorithms related to the open problems listed in topics of interest and **(b) position/vision** contributions discussing uses cases, challenges, and future research directions. These contributions should be at least 4 pages, but up to 12 (excluding references), formatted in CEURART's single-column template. Accepted contributions will be presented at the workshop and published in workshop proceedings.

Editorially-reviewed. To enable potential attendees to articulate their views on the topics of the workshop, share

[3]https://sigir2025.dei.unipd.it/low-resource-environments-track.html

already-published works, bring awareness to ongoing European projects in this area, etc., we will also elicit 1 page-extended abstracts. This type of informal submission will be editorially reviewed by workshop organizers to gauge fit as well as themes of interest for group discussions. Accepted contributions will be presented at the workshop, but not published in the proceedings.

The workshop will be structured to foster interaction among participants. To facilitate networking and interaction among attendees, and in the spirit of building community we will start the workshop with a brief welcome presentation and ice-break activity to encourage workshop participants to share their background, point of view and interest in IR4U2. This will be followed by a 'Fireside Chat' keynote talk, moderated by one of the workshop organizers. The purpose of this activity is to depend upon an expert in the user-centered AI arena to prompt a discussion that will inspire and inform the discussion planned for the rest of the day. Presentations of accepted contributions will be brief to keep the workshop's flow vibrant and to allow time for focused group work.

We will allocate time to work in smaller groups so that all attendees can have a voice and share their perspectives. We anticipate group discussion as emerging from the analysis of intention forms and submissions to the workshop. Activities aiming to facilitate discussion include sticky notes and large shared paper techniques. Following the previous phase, we will facilitate a collaborative brainstorming session where all groups come together to identify common algorithmic biases, shared ranking challenges, and potential modifications to IR models, relevance metrics, and interaction models.

Up-to-date information on the workshop program and accepted contributions can be found on the dedicated IR4U2 website[4].

3 Workshop Organizers

Below, we include a summary of the organizer's background.

Noemi Mauro is a Tenure-track Assistant Professor at the Computer Science Department of the University of Turin where she obtained a PhD in Computer Science with Honors. Her research interests concern user modeling, recommender systems, cultural heritage, information filtering, and information visualization. She is a PC member of the top conferences in her research areas and a reviewer for several related journals. She co-edited the special issue "Intelligent Systems for People with Diverse Cognitive Abilities" in the Human-computer Interaction journal. She is an Editorial Board Member of the User Modeling and User-Adapted Interaction journal.

Angelo Geninatti Cossatin is a PhD candidate at the Computer Science Department of the University of Turin. His research interests concern recommender systems, user modeling and human-computer interaction. He is currently working on projects related to understudied users, namely people with ASD.

Maria Soledad Pera is an Associate Professor at the Web Information Systems group of the EEMCS faculty at TU Delft. Sole's research focuses on Information Retrieval, with a special emphasis on enhancing information access for underserved user groups. She serves as (S)PC for SIGIR, UMAP, CHIIR, RecSys, and ECIR. She was General Chair for RecSys '18 and Program Chair for UMAP '23 and CIKM '24. Among others, she co-organized the ComplexRec workshop and 6 editions of the KidRec workshop (on search and recommendation for children).

Federica Cena is an Associate Professor at the Computer Science Department of the University of Turin. She works on the intersection of Artificial Intelligence and Human-Computer Interaction. Her recent research has studied the implications of the Internet of Things for user modeling and personalization, with a special focus on assistive applications for cognitive disabilities and frailty.

Theo Huibers has been researching information retrieval and human media interaction for over 30 years. Since 2002, he has been a professor in Human Media Interaction & Computer Science at the University of Twente, focusing on IR for children.

Monica Landoni is a titular professor at the faculty of Informatics at Università della Svizzera Italiana (USI). She is Chair of the ACM Interaction Design and Children Conference Steering Committee and member of EUGAIN, the European Network For Gender Balance in Informatics. She has worked on national and European projects investigating how technology can support children when searching, writing, and reading for education and pleasure.

Emiliana Murgia With a degree in Literature and courses, including Communication and Management of School and Training Institutions, Emiliana transitioned from a career in communication to teaching at the primary school in 1999/2000; actively promoting technology in education, she has been a "digital animator" at school, managing various national and international projects, and has collaborated with the University of Milano Bicocca since 2014 on teaching with technologies. In 2018, she joined a multidisciplinary research team focusing on online information access, including AI. Currently on secondment to participate in a National Doctorate program in Learning Sciences and Digital Technologies.

[4]https://ir4u2workshop.wixsite.com/ir4u2

References

[1] Ayana Archie. 2023. A U.K. agency has fined TikTok nearly $16 million for handling of children's Data. https://www.npr.org/2023/04/05/1168114842/tiktok-uk-fine

[2] Ion Madrazo Azpiazu, Nevena Dragovic, Maria Soledad Pera, and Jerry Alan Fails. 2017. Online searching and learning: YUM and other search tools for children and teachers. *Information Retrieval Journal* 20 (2017), 524–545.

[3] Judith H Danovitch. 2019. Growing up with Google: How children's understanding and use of internet-based devices relates to cognitive development. *Human Behavior and Emerging Technologies* 1, 2 (2019), 81–90.

[4] Pablo Delgado, Vicenta Ávila, Inmaculada Fajardo, and Ladislao Salmerón. 2019. Training young adults with intellectual disability to read critically on the internet. *Journal of Applied Research in Intellectual Disabilities* 32, 3 (2019), 666–677.

[5] Christopher D Manning. 2009. *An introduction to information retrieval*. Cambridge university press.

[6] Noemi Mauro, Liliana Ardissono, and Federica Cena. 2022. Supporting People with Autism Spectrum Disorders in the Exploration of PoIs: An Inclusive Recommender System. *Commun. ACM* 65, 2, 101–109. https://doi.org/10.1145/3505267

[7] Guy Meyer, Alan Wassyng, Mark Lawford, Kourosh Sabri, and Shahram Shirani. 2022. Literature Review of Computer Tools for the Visually Impaired: A Focus on Search Engines. *Artificial Intelligence in Healthcare and Medicine* (2022), 237–259.

[8] Ashlee Milton and Maria Soledad Pera. 2023. Into the Unknown: Exploration of Search Engines' Responses to Users with Depression and Anxiety. *ACM Transactions on the Web* 17, 4 (2023), 1–29.
[9] Safiya Umoja Noble. 2018. Algorithms of oppression. In *Algorithms of oppression*. New York university press.
[10] Maria Soledad Pera, Federica Cena, Monica Landoni, Cataldo Musto, and Alain D Starke. 2024. Human Factors in User Modeling for Intelligent Systems. In *A Human-Centered Perspective of Intelligent Personalized Environments and Systems*. Springer, 3–42.
[11] Konstantinos Tsiakas, Emilia Barakova, Javed Vassilis Khan, and Panos Markopoulos. 2020. BrainHood: towards an explainable recommendation system for self-regulated cognitive training in children. In *Proc. of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 1–6.
[12] Suzan Verberne, Evangelos Kanoulas, Gineke Wiggers, Florina Piroi, and Arjen P de Vries. 2023. ECIR 2023 Workshop: Legal Information Retrieval. In *European Conference on Information Retrieval*. Springer, 412–419.
[13] Ryen W White, Elad Yom-Tov, Eric Horvitz, Eugene Agichtein, and William Hersh. 2013. Report on the SIGIR 2013 workshop on health search and discovery. In *ACM SIGIR Forum*, Vol. 47. ACM New York, NY, USA, 101–108.
[14] Bo Xie, Neil Charness, Karen Fingerman, Jeffrey Kaye, Miyong T Kim, and Anjum Khurshid. 2021. When going digital becomes a necessity: Ensuring older adults' needs for information, services, and social inclusion during COVID-19. In *Older Adults and COVID-19*. Routledge, 181–191.