Report on the 1st IR for Children 2000-2020: Where Are We Now? (IR4C) Workshop at SIGIR 2021: The Need to Spotlight Research on Children Information Retrieval

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

Emiliana Murgia
University of Milano Bicocca
emiliana.murgia@unimib.it

Theo Huibers
University of Twente – Wizenoze
t.w.c.huibers@utwente.nl

Maria Soledad Pera
PIReT- Boise State University
solepera@boisestate.edu

Abstract

We present a brief overview of the discussions and takeaways emerging from the IR for Children 2000-2020: Where Are We Now? Workshop, co-located with the 44^{th} International ACM SIGIR Conference on Research and Development in Information Retrieval.

Date: 15 July, 2021.

Website: http://www.fab4.science/IR4C/.

1 Introduction

Children (individuals under the age of eighteen[1] regularly turn to information retrieval tools, such as search and recommendation systems, for information access. Yet, this particular user group–with a broad range of skills and expectations–is seldom the focus of research endeavors in the realm of Information Retrieval (**IR**).

As reported in [Huibers et al., 2021a; White, 2021], over the past two decades IR researchers and industry practitioners have turned their attention to this audience, as evidenced by efforts like Yahooligans![2], PuppyIR [Azzopardi et al., 2009], workshops like *Towards Accessible Search Systems* [Serdyukov et al., 2011] or *KidRec* [Huibers et al., 2021b], in addition to research works that sporadically make their way into core IR-related conferences. Unfortunately, children continue to struggle interacting, for example, with commercial search engines when trying to satisfy their information needs and lack information access literacy, which is a must in an era of misinformation). Further, in domains like education, we see a lack of consensus on what technology

[1] https://www.ohchr.org/en/professionalinterest/pages/crc.aspx
[2] https://en.wikipedia.org/wiki/Yahoo!_Kids

should do to balance engagement and support of students and the expert in the loop (the teachers). These are some of the issues that steered us to propose the 1^{st} edition of the IR for Children 2000-2020: Where Are We Now? (**IR4C**), a workshop that took place during the 2021 ACM SIGIR Conference on Research and Development in Information Retrieval.

The goal of IR4C was to assemble researchers and practitioners who, over the last two decades years, have directly or indirectly contributed to the Children Information Retrieval (**CIR**) community. In particular, we aimed to foster a discussion among attendees that could help us take inventory of where we are as a community when it comes to CIR while suggesting directions as to where we want and should go.

In the rest of this report, we revisit the outcomes from IR4C. We first briefly describe the workshop program, the keynote address, and the accepted contributions. Then, we discuss the main takeaways, concluding remarks, and future directions for the CIR community.

2 Workshop Organization

We summarize below IR4C's program, in addition to the keynote address, the presented contributions, and issues and themes emerging from group discussions.

2.1 Program

Given the restrictions imposed by the ongoing COVID-19 pandemic, we hosted a half-day workshop over Zoom[3]. By allocating only half a day, we accommodated participants from around the world (including from New Zealand, the US, Africa, and Europe) while avoiding the fatigue commonly associated with online work gatherings [Nadler, 2020].

We were honored to have Professor Bilal as a keynote speaker. Her overview of the history of CIR research and open paths for the future, in addition to the presentations by authors from accepted contributions, sparked stimulating discussions among workshop attendees. Given that the workshop took place remotely, to facilitate a highly participatory workshop[4], one promoting interactions and exchanges of ideas, we encouraged workshop participants to take notes on research questions and challenges inspired by the invited speaker and presenters throughout the day using Miro[5].

2.2 Keynote Address

We started the day with a keynote address by Professor Dania Bilal (School of Information Sciences – University of Tennessee-Knoxville).

Abstract. Research on children's information interaction and retrieval has flourished in the past two decades. The rise of the Google search engine in the late 1990s revolutionized how people access, search, and retrieve information on the internet; its simple design and popularity have influenced the development of many IR systems. For children, Google remains the most preferred

[3] https://zoom.us

[4] https://we.riseup.net/assets/25682/FacilitatingWorkshops.pdf

[5] https://miro.com

search engine, and its increased use in and outside of school has piqued interest among researchers in investigating children's cognitive and emotional patterns in interacting with the engine. For example, in 2010 the SIGIR Workshop on Accessible Search Systems featured state-of-the-art research on children's information interaction with different search systems. In a similar vein, the ubiquitous use of social media among children has driven researchers to explore children's information interaction in various media platforms, unveiling emotional, social, and information literacy concerns, as well as design issues. Most recently, there has been a focus on interaction with conversational Artificial Intelligence (AI), such as chatbots and voice digital assistants. Nonetheless, the body of research on children's information interaction with IR systems is relatively small, compared to that in the information retrieval and human information behavior fields.

While the past decade brought strides in the design of human-centered IR systems accessible to children (e.g., auto-complete or predictive text, query suggestions, spelling corrections, relevance ranking, suggestions for related or similar searches, and rich snippets), much work is still needed.

The introduction of AI-driven solutions (e.g., robots; drones) in school and the increasing usage of smartphones among people of all age groups, including children, could lead to a new paradigm in children's information interaction research and theory in the next few years. Such a paradigm might necessitate ideations among scholars and system designers to develop creative and innovative solutions for designing adaptable child-centered IR systems.

Discussion. Professor Dania Bilal shared her perspectives on the growth of the field of research focused on studying children's interactions with IR systems. She also addressed issues and challenges in the current research landscape and outlined hurdles and opportunities inherent to the design of IR systems for children. More importantly, Dr. Bilal draw attention to *where we are* and *where we should be* in the next decade as a community of practice. One of the statements that most resonated with attendees and prompted conversation later on during the workshop was that literature focused on CIR does leverage knowledge emerging from research endeavors focused on IR for mainstream individuals (in their majority adults). The opposite, however, does not hold. Dr. Bilal encouraged IR4C attendees to help bring awareness among the IR community at large on the importance to recognize the implications of research done when children are the main stakeholders [White, 2021]. This is because valuable findings from work related to CIR could serve as another layer towards advancing knowledge in mainstream IR.

2.3 Accepted Contributions

In the call for papers for IR4C, we invited vision papers sharing past experiences, ongoing projects, and future directions referring to IR systems for whom children are the primary stakeholders. Overall, we accepted six contributions, which covered a broad range of open problems aligning with CIR. We invited authors from these accepted submissions to give a short presentation.

Each submission was peer-reviewed by at least three program committee (PC) members, consisting of international experts. We want to thank the PC members who selflessly dedicated time to ensure the quality and potential to spark discussion of the accepted contributions. In addition to IR4C workshop organizers, PC included: Mohammad Aliannejadi (University of Amsterdam), Katherine Landau Wright (Boise State University), and Ian Ruthven (University of Strathclyde).

Vanderschantz and Hinze [2021] presented two models of children's internet search behavior in an educational context–one capturing children's perceptions and another capturing their teachers'.

They juxtaposed these models with a researcher's view on how to depict children's search behavior in a classroom setting. Contrasting these three models enabled the authors to outline several areas that need attention, including conducting more in-depth in-situ studies with educationally appropriate study parameters set by or in consultation with teachers. Moreover, Vanderschantz and Hinze [2021] argue that "while current adult IR tools and search engines do not meet the needs of children, we believe that IR systems specifically designed for children are limiting". The presentation by Nicol [2021] also examined the educational context. She shared her experiences of studying classroom technology from the IR and human-computer interaction perspectives. In particular, she mentioned potential limitations of the current understanding of information seeking behavior, as it might be too reliant on data from studies conducted in controlled (i.e., synthetic) environments. Among future challenges for IR researchers to consider, the author indicated the need to move beyond text when it comes to search technology to instead consider voice-activated devices; with the caveat of taking into account the logistical and environmental implications of a class of up-to-30 students concurrently using these devices.

Valguarnera [2021] brought to our attention the role parents and teachers should have when it comes to the design, implementation, and adoption of search tools for very young children, ages 3 to 6 years old. The author discussed the prominence of voice interfaces for search, yet the difficulty of incorporating such interfaces in a classroom setting or the preferences of parents or teachers regarding the purpose of letting children engage with said interfaces. She also urged researchers to engage in co-design activities to produce search tools targeting the needs of these very young searchers, if possible with the support of parents as informants.

The rest of the presentations tackled three concepts inherent to technology enabling children's information access, regardless of the context (for learning or leisure): trust [Westerveld et al., 2021], accountability [Beelen et al., 2021], and accessibility [Milton et al., 2021]. Westerveld et al. [2021] presented technology developed by Wizenoze to help learners in primary, secondary, and vocational education find the best online resources. The authors suggested that key factors guiding the design of technology in this context include the complexity of resources, their suitability and recency, curriculum requirements, in addition to how resources align with these requirements as factors impacting what makes a resource trusted. Beelen et al. [2021] brought up challenges related to CIR and identified conversational robots as a means to ease this process responsibly. The authors pointed out that tools children currently use, such as search engines, do not always meet their specific needs. With that in mind, they proposed developing a conversational robot that could maintain context, ask clarifying questions, and give suggestions that better meet children's needs. As children tend to be far too trusting of robots, the authors emphasize the importance for the robot to measure, monitor, and adapt to the level of trust of each child user. Lastly, Milton et al. [2021] reminded us how seminal works have studied how children use mainstream search engines, as well as how to design and evaluate custom search engines explicitly for children–often treating children as a unit. Even at the same age, however, children are known to possess and exhibit different capabilities. These differences affect how children access and use search engines. This motivated the authors to focus on the concept of accessibility and discuss why current research on children and search engines does not, but should, focus on this significant matter.

2.4 Group Discussions

We invited participants to use an online collaborative whiteboard platform (i.e., Miro) to post observations, reflections, open issues, and concerns as triggered by the presentations described in Sections 2.2 and 2.3 to stimulate group discussion. For the same purpose, we asked workshop attendees to explicitly list research questions and challenges about CIR.

It became apparent that a few of the contributions drove attendees to focus on the reasons behind children's favoring Google over custom-made search tools. A timely by one of the participants increased attendees' perceptions on how children change faster than the IR systems. Indeed, researchers and practitioners must be aware that they are now dealing with the TikTok generation. There were concerns about children's lack of search literacy, with suggestions on how to make search tools more child-friendly by providing needed scaffolding. Other surfacing open questions spanned from how to conciliate the need for personalization with privacy to how to define what it means for online resources provided by IR systems to be 'good'. Additional areas of interest connected how the digital divide could become more severe as the sum of the access and expertise of the involved adults (parents and teachers) with that of children and their level of familiarity and proficiency with technology. Moreover, workshop participants questioned the feasibility (and mentioned the associated challenges) of aligning educational curricula with search literacy, let alone establishing a TREC-like initiative to help CIR grow.

As a result of clustering Miro board sticky notes produced by participants throughout the workshop and further examining group discussion, we detected some developing trends. More prominent is the quest for user studies to be run across countries, which would enable the CIR community to produce a deeper understanding of the needs of children in regards to technology that supports information access while taking into account different languages, curricula, and cultures. In turn, this would allow us to advance knowledge on how and what we can learn from specific studies to support 'all' children. The design of such a study is non-trivial. It requires that the CIR community outlines several basic concepts, beginning with what would be a suitable task to serve as the focus of this study and who can define it, how to recognize a good resource, and which traits best describe an independent child searcher. Other essential areas that attendees gravitated towards included the importance of children being actively involved in the design of novel IR tools, the necessity to incorporate teachers' and parents' voices in the design process, and the potential of novel interfaces (e.g., wearables). A specific use case in this area is a multimedia search tool that could be mediated somehow by search agents in the form of robots or chatbots, and the role these agents can play in the classroom. Finally, it is worth emphasizing the merit of reflecting on how much we have learned from studies involving adults compared with the limited attention that has been given to CIR by researchers focusing on how adults search.

3 Concluding Remarks and Next Steps

Using the question where are we now? as an initial point for reflection, we can conclude after the workshop that we have come far: over the past twenty years, a great deal of research has taken place in various sub-areas, from child-media interaction to media literacy. Unfortunately, we must also acknowledge that there are many–really many–research questions that remain unanswered.

We posit that there are two main reasons why CIR research is not advancing at the pace of

core IR research:

1. In her speech, Dr. Bilal mentioned that researchers and designers often see children as 'lesser adults'. Then, an easy solution for the CIR problem involves simply trying to 'downsize' general IR systems. Nothing is further from the truth. Perhaps children are the purest users of computer systems, and they use IR systems freely.
2. As a research area of study, CIR is not integrated. Each sub-area offers its own perspective, i.e., each sub-area yields solutions that serve as a piece of a puzzle, but the big puzzle is never put together.

The enthusiasm of our participants and their lively conversations are testimony of the importance and the interest for CIR. As workshop organizers–who also happen to be part of an interdisciplinary team, denoted FAB4[6], consisting of researchers and practitioners interested in advancing knowledge in CIR–we feel compelled to keep the momentum going. This can be achieved by offering a more frequent channel for researchers to meet and exchange concerns and proposals for future collaborations. Consequently, we are working on the design of an initial cross-country usage study for eliciting data on how children use search tools, which at the same time could allow us to better understand which factors influence children's search behavior. We envision this study as a first step towards the design and international deployment of a questionnaire inquiring on the use, requirements, and expectations related to IR in the classroom. The design of such a study will give our community a shared intend and a practical exercise to engage with. We hope it will also help us establish essential terminology, define common practices, explore innovative interfaces, and, more importantly, stimulate discussion on ethical issues. We also hope this joint effort will attract the attention of the mainstream IR community and open up to a more balanced and mutually beneficial exchange of results and achievements.

We try to continue to advance knowledge related to CIR, with as many people as possible from different disciplines and of course with children and teachers included! If you would like to broaden and strengthen our working group, please mail us (info@fab4.science) and we will keep you informed about future initiatives.

References

Leif Azzopardi, Richard Glassey, Mounia Lalmas, Tamara Polajnar, and Ian Ruthven. Puppyir: Designing an open source framework for interactive information services for children. In *Proceedings of the Annual Workshop on Human-Computer Interaction and Information Retrieval*, volume 44, 2009.

Thomas Beelen, Ella Velner, Roeland Ordelman, Khiet Truong, Vanessa Evers, and Theo Huibers. Does your robot know? enhancing children's information retrieval through spoken conversation with responsible robots. In *Proceedings of the IR for Children 2000-2020: Where Are We Now?, workshop, co-located with ACM SIGIR*, 2021. URL https://arxiv.org/abs/2106.07931.

Theo Huibers, Monica Landoni, Emiliana Murgia, and Maria Soledad Pera. Ir for children 2000-

[6] https://www.fab4.science/

2020: Where are we now? In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2689–2692, 2021a.

Theo Huibers, Monica Landoni, Maria Soledad Pera, Jerry Alan Fails, Emiliana Murgia, and Natalia Kucirkova. What does good look like? report on the 3rd international and interdisciplinary perspectives on children & recommender and information retrieval systems (kidrec) at idc 2019. In *ACM SIGIR Forum*, volume 53, pages 76–81. ACM New York, NY, USA, 2021b.

Ashlee Milton, Garrett Allen, and Maria Soledad Pera. To infinity and beyond! accessibility is the future for kids' search engines. In *IR for Children 2000-2020: Where Are We Now? Workshop co-located with ACM SIGIR*, 2021. URL https://arxiv.org/abs/2106.07813.

Robby Nadler. Understanding "zoom fatigue": Theorizing spatial dynamics as third skins in computer-mediated communication. *Computers and Composition*, 58:102613, 2020.

Emma Nicol. Information retrieval for children in educational contexts: Can we keep up? In *IR for Children 2000-2020: Where Are We Now? Workshop co-located with ACM SIGIR*, 2021. URL https://drive.google.com/file/d/19LLVpyQHbijdWBsIr_Q3_0O-z-qOL5p1/view.

Pavel Serdyukov, Djoerd Hiemstra, and Ian Ruthven. Towards accessible search systems. In *ACM SIGIR Forum*, volume 44, pages 23–27. ACM New York, NY, USA, 2011.

Sveva Valguarnera. The role of parents and preschool teachers in the design, implementation and adoption of search tools for young children. In *IR for Children 2000-2020: Where Are We Now? Workshop co-located with ACM SIGIR*, 2021. URL https://drive.google.com/file/d/1AVYBhj8YvetprTdsLeXJZtVGvRhskAfe/view.

Nicholas Vanderschantz and Annika Hinze. Children's internet search behaviour in an educational context. In *IR for Children 2000-2020: Where Are We Now? Workshop co-located with ACM SIGIR*, 2021. URL https://drive.google.com/file/d/1UfSD6g3TxZEAU7YiaM_jGmbXkB1qm9JT/view.

Thijs Westerveld, Gerben De Vries, Hanna Jochmann-Mannak, Carsten Schnober, Chrisine Fraser, and Theo Huibers. Trusted digital content for learners. In *IR for Children 2000-2020: Where Are We Now? Workshop co-located with ACM SIGIR '21*, 2021. URL https://drive.google.com/file/d/1AYBC9W4Zk5HarpbEVrtB1G82qkecJxgS/view.

Martin White. Report on the sigir 2021 workshop "ir for children 2000-2020: Where are we now?", Aug 2021. URL https://irsg.bcs.org/informer/2021/08/report-on-the-sigir-2021-workshop-ir-for-children-2000-2020-where-are-we-now/.