

4th KidRec – What does Good Look Like: From Design, Research, and Practice to Policy

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

Maria Soledad Pera
People and Information Research Team
Boise State University – Boise, ID, USA
solepera@boisestate.edu

Jerry Alan Fails
Dept. of Computer Science
Boise State University – Boise, ID, USA
jerryfails@boisestate.edu

Emiliana Murgia
Università di Milano-Bicocca
Milano, Italy
emiliana.murgia@unimib.it

Natalia Kucirkova
University of Stavanger
Stavanger, Norway
natalia.kucirkova@uis.no

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

Abstract

We summarize outcomes from KidRec 2020, the 4th edition of the workshop on International and Interdisciplinary Perspectives on Children & Recommender and Information Retrieval Systems; which this year was co-located with the ACM Interaction Design and Children conference. In addition to lessons learned from the keynote address by Professor Marcia C. Linn, accepted contributions, and collaborative work during the workshop, we offer insights related to hosting KidRec as a virtual meeting during the COVID-19 pandemic.

1 Introduction

Since 2017, when we held the first iteration of the KidRec workshop (co-located with the ACM Recommender Systems conference), we have continued to build community around an important and rich topic: the design, development and evaluation of Information Retrieval (IR) systems – search and recommendation systems – that explicitly target children.

We first focused on *recommender systems* [Pera et al., 2018]; then, we expanded to IR systems in general and shined a light on those that are meant to support *education* [Fails et al., 2019]. Last year, we focused on how to *evaluate* the IR systems that are tailored to users [Huibers et al., 2019]. While outcomes from that 3rd edition of KidRec revealed preliminary insights into what makes an IR system for children “good”, as a community we noticed that there are many questions that

remain unanswered. With that in mind, we decided that the theme of the 4th edition of KidRec would help us continue our inquiry into what we mean by good related to IR systems for children; this time the statement driving our discussion forward was *What does good look like: from design, research, and practice to policy* [Landoni et al., 2020].

In the rest of this manuscript, we offer an overview of the keynote address, the accepted contributions, and an overview of the discussions from the group interactions directed towards continuing to build community around this important area of study. We also discuss the next iteration of our evaluation framework, and offer insights pertaining to the fact that this edition of KidRec was hosted online due to the COVID-19 pandemic.

2 Keynote

We were fortunate to have Dr. Marcia Linn, Professor of Development and Cognition, who specializes in science and technology at the Graduate School of Education, University of California, Berkeley, start the conversation with her keynote.

Bio. Dr. Linn is a member of the National Academy of Education and a Fellow of the American Association for the Advancement of Science (AAAS), the American Psychological Association, and the Association for Psychological Science. She has served as President of the International Society of the Learning Sciences, Chair of the AAAS Education Section, and on the boards of the AAAS, the Educational Testing Service Graduate Record Examination, the McDonnell Foundation Cognitive Studies in Education Practice, and the National Science Foundation. Awards include the National Association for Research in Science Teaching Award for Lifelong Distinguished Contributions to Science Education, the American Educational Research Association Willystine Goodsell Award, and the Council of Scientific Society Presidents first award for Excellence in Educational Research. Linn earned her Ph.D. at Stanford University where she worked with Lee Cronbach. She spent a year in Geneva working with Jean Piaget. She has been a fellow at the Center for Advanced Study in Behavioral Sciences three times. Her books include *Internet Environments for Science Education* (2004), *WISE Science* (2009), and *Science Teaching and Learning: Taking Advantage of Technology to Promote Knowledge Integration* (2011). She chairs the *Technology, Education—Connections* series for Teachers College Press.

Abstract. This talk discusses ways we are using real-time, autoscored student data along with research-derived instructional customizations via a digital teacher report (the Teacher Action Plan, or TAP) to inform both teacher practice (implementation of responsive instruction) and student learning (as evidenced by their written and revised explanations for an embedded milestone assessment item).

Discussion. Dr. Linn presented an overview of STRIDES, a project focused on responding to three areas of requirements defined in the Next Generation Science Standards: (i) core ideas, (ii) practices, and (iii) crosscutting. She then went on to discuss the development of a knowledge integration framework she developed, that examines how learners integrate or synthesize multiple models together into a unified model. Lastly, she introduced Web-based Inquiry Science Environment¹, a project meant to offer online learning, evaluation of responses, and scaffolding to support teachers provide informative feedback.

¹WISE project: <https://wise.berkeley.edu/>

During the Q&A session, participants were able to connect concepts presented by Dr. Linn to their own varied research and perspectives, including the retrieval of information, spelling correction, and education aided by technology, to name a few.

3 Interactive Panel I - Accepted Contributions

During the first interactive panel session, we observed brief presentations of the accepted paper contributions, followed by a discussion of the presented materials. The abstracts of the accepted contributions are below, followed by a synopsis of the discussion.

What does good look like when working in the classroom with 3-6 years old children? by Sveva Valguarnera (Università della Svizzera Italiana)². Working with preschool children in their classrooms present unique challenges, due to the specific age range and the characteristics of the environment. In this position paper, I will detail the challenges that I have encountered during a user study conducted in a Swiss preschool.

#TheHorror: Evaluating Information Retrieval Systems for Kids by Ashlee Milton (Boise State University) [Milton and Pera, 2020]. Evaluation of information retrieval systems (IRS) is a prominent topic among IR researchers—mainly directed at a general population. Children require unique IRS and by extension different ways to evaluate these systems, but as a large population that use IRS have largely been ignored on the evaluation front. In this position paper, we explore many perspectives that must be considered when evaluating IRS; we specially discuss problems faced by researchers who work with children IRS, including lack of evaluation frameworks, limitations of data, and lack of user judgment understanding.

KidRec 2020: Position Paper by Emma Nicol (University of Dundee)³. My recently completed PhD thesis (2019) describes a school-based study that investigated the information seeking behaviour of primary school children in Scotland. In this presentation, I will share lessons learned from this thesis, along with other past experiences, as I have been involved in several projects investigating children’s search behaviour and the design and use of technology with and for children e.g., WEBKIT on tangible search technology for children (2004) and Wi Gleesome Touch on museum interactive games (2010).

Discussion. The common denominator among the presentations were the challenges associated with conducting user studies/evaluations involving children – specially when studies involved the use of new technologies. Across the presentations, a question continued to arise: how do children perceive relevance; what does success mean to them? This could vary from: aesthetic characteristics, to personal preferences, to the quality (in some dimension, e.g.: educational value, readability) of the resources recommended or presented as a result of a search they would use. It was clear that the concept of *relevance* for children goes beyond the more traditional perspective of query-document mapping [Crestani et al., 1998].

Another interesting discussion related to the discrepancies between children’s and adults’ perspectives: presenters revealed that while some parents considered their children non-experts in the use of technology, children were quick to use and interact with technology. The child’s current temperament can also hinder study completion. In fact, shyness was one of the reasons cited as

²Available at: https://kidrec.github.io/papers/KidRec_2020_paper_1.pdf

³Available at: https://kidrec.github.io/papers/KidRec_2020_paper_3.pdf

to why children were uncomfortable with some studies conducted by workshop participants. Familiarity with research and having participated in previous user studies also influences outcomes; children who experienced previous study sessions were much more willing to engage and actively participate in studies.

From the presentations it also became apparent that evaluation enabling comparison across strategies remains an open problem for search and recommendation systems targeting young audiences, as data-sets that would enable such an analysis are not publicly available. Further, data remains a concern: dealing with a protected population makes it so that data collection and sharing are highly sensitive, as they should be, and therefore need to be handled with care and caution. Yet, that sensitivity can hinder design and development of more effective IR systems that better serve young users.

Last, but not least, a compelling point was made: findings and results reported in the literature pertaining to children’s search tasks at school are not, for the most part, based on authentic (e.g. *in situ*) studies. Instead, they are based on tasks created and proposed by researchers. This evidences the need for authentic user studies that enable better capturing children’s information seeking process: their expectations and challenges.

4 Interactive Panel II - Participant Perspectives

This year, we also invited participants without an accepted contribution to briefly share their perspectives and expertise related to the theme of the workshop.

Researchers from the European Commission Joint Research Centre – Institute for the Protection and Security of the Citizen, shined a light on the importance of *privacy*. They paid special attention to the need to consider the privacy paradox in order to ensure safety and security of IR systems, given that children are the main stakeholder. This paradox describes the disparity between people’s privacy intentions and their behaviors [Norberg et al., 2007].

We also heard from young researchers involved in the design of a child-adaptive search tool (CAST)⁴. Particularly, they shared challenges so far on *data collection*, given they are dealing with a protected audience. The lack of data is a major issue as data is needed in order to evaluate and improve outcomes.

From an interdisciplinary group of researchers, we learned about the *binary* view children have when deciding resource relevance in response to queries submitted to popular search tools. From preliminary results it emerged that among results perceived as relevant, their position in the ranking did not entice children to consider top-ranked one as more relevant than lower-ranked counterparts. This calls for revisiting evaluation frameworks when it comes to IR systems for children and make us wonder which metrics could suitably address such concern.

5 Building the Framework

In the last part of the workshop, we divided into small groups to discuss the next iteration of our framework. We started with the four pillars of a “good” IR system for children that we established

⁴Details on CAST search tool can be found at: <https://cast.boisestate.edu/>

at last year’s KidRec. An IR system for children is good if: (1) It provides **resources** that are logically relevant, useful, and foster learning; (2) It is designed with a **user-centered** perspective while acknowledging that multiple stakeholder perspectives need to co-exist; (3) It is **ethically** sound and supports children’s rights; and (4) Users are deeply **engaged** with the system.

Thereafter, we tasked each of the small groups to discuss and propose ideas on how to take each of the pillars from theory to practice. To foster interaction, we had teams create (virtual) sticky-notes briefly describing their ideas related to practical applications that could help quantify each of the aforementioned pillars.

The small groups struggled to move the theory to practice and came up with additional challenges which we noted as representative of the complex landscape within this area. Asynchronously, workshop organizers continued to work on merging and labeling different groups of ideas. This resulted in several themes and outcomes. Additional challenges were clustered into the themes of: human elements, system requirements, ethical issues, and environmental factors. With regards to human elements, questions were raised on “how do we engage children”, who are all of the relevant stakeholders, the need to carefully incorporate the various stakeholder perspectives (parents, teachers, in addition to children), how do you meld and give relative importance to the evaluation of each stakeholder’s perspective, and how to correlate or accommodate age and developmental stage within the evaluation framework. In terms of system requirements, in addition to reliability and ease of access, there was a concern for increasing inclusivity – including but not limited to working on older hardware and being accessible to a broad spectrum of users with varied abilities and needs. Ethical issues that emerged included protecting children, giving children an authentic voice, addressing power dynamics so that adults don’t just “take control”, avoiding bias, and assuring proper ethical review processes (e.g., review boards and informed consent) are followed. The last cluster related to tasks, environment, and how to evaluate those tasks. For tasks, concerns included real (authentic) versus artificial tasks and browsing versus searching. Environment or context of use also has a big impact as well – whether the IR system is being used in a school, at home, or on the move. When it comes to evaluation, common concerns included what is the ground truth, what data sources are available, how to evaluate resources as relevant and/or useful for learning, what are the shades or gradation of relevance, and what are the appropriate metrics for evaluation (beyond just precision and recall).

In considering these and previously identified challenges, the group worked to continue to discuss how to take the pillars identified at the 3rd edition of KidRec from theory to practise, in the pursuit of a framework for design and evaluation of IR system for children. From discussions, it became apparent that in next iterations of the framework:

1. **Resources** should cover available models of learning objects, suitable content including open access material for education and multimedia material specific for children, while also include metrics, measures and standards to guide their evaluation.
2. **User-centered** should recognize and stress how each child is unique and the importance of catering for diversity with specific attention to the roles played by multiple stakeholders, e.g. teachers, parents, siblings, industry, designers and developers while accounting for their different cultural backgrounds and philosophical beliefs.
3. **Ethical** concerns should deal with the tension between giving a voice to children and protecting their identity, with emphasis on the importance of child friendly informed consent

as a tool to ensure respect beyond the legal requirements.

4. **Engagement** is a crucial criterion, linked with time spent in searching and exploring. Interest in task and motivation for finding good results are also important factors contributing to the level of engagement with the IR tool.

6 Virtual Mode: On

The main challenge of hosting a workshop remotely was to foster interactions among participants early on so as to ensure effective group work throughout the workshop and prompt networking opportunities beyond. With that in mind, and taking advantage of Breakout Room functionality offered by Zoom⁵ we divided attendees into small groups, and had them introduce themselves. Once all attendees merged back into Zoom’s main room, we asked each participant to introduce their fellow participant from the group. The smaller gatherings made it so that each attendee had at least one other counterpart to “navigate” the workshop and share breaks with.

While we originally envisioned KidRec 2020 as a full-day workshop, we soon realized that it was not feasible when conducted as an online workshop, due to the different time-zones of all participants, and more importantly, due to the now known effect of “Zoom Fatigue” [Fosslien and Duffy, 2020; Jiang, 2020]. This is why we reorganized workshop “live” activities to last for at most 3.5 hours. On the upside, workshop attendees were highly active throughout the duration. On the downside, anticipated group work related to evaluation framework iteration (see Section 5) could not be completed. This is what prompted us to employ asynchronous activities that continued for two weeks beyond the workshop, using Slack⁶ as a means of communication and Miro⁷ as platform for collaborative work. This enabled us as workshop organizers to achieve our intended goals of creating a new iteration for an evaluation framework that allows assessment and comparisons across IR systems for children while fostering an inclusive and interactive work environment that welcomes all varied stakeholders involved in this area of study, including students, researchers, and practitioners in computer science, education, and other adjacent domains.

7 Conclusions and Next Steps

We continue to build a community that addresses important issues for children and recommender systems. From the interactive discussions that we conducted, it became clear that more work needs to be done in the area in order to bring theoretical concerns into a practical framework that researchers and practitioners from varied areas of research can turn to when it comes to designing and evaluating IR systems tailored towards children.

Mindful of the ongoing and new challenges that were discussed in this 4th edition of KidRec, we are currently developing a call for papers for a special issue of International Journal of Child-Computer Interaction, focused on design and evaluation of IR systems for children. Authors of

⁵<https://zoom.us/>

⁶<https://slack.com/>

⁷<https://miro.com/index/>

accepted contributions to the special issue will be invited to present their work at the KidRec workshop in 2021.

References

Fabio Crestani, Mounia Lalmas, Cornelis J. Van Rijsbergen, and Iain Campbell. “is this document relevant?... probably”: A survey of probabilistic models in information retrieval. *ACM Comput. Surv.*, 30(4):528–552, December 1998. ISSN 0360-0300.

Jerry Alan Fails, Maria Soledad Pera, and Natalia Kucirkova. Building community: Report on the 2nd international and interdisciplinary perspectives on children & recommender systems (kidrec) at idc 2018. In *ACM SIGIR Forum*, volume 52, pages 138–144. ACM New York, NY, USA, 2019.

L Fosslien and M West Duffy. How to combat zoom fatigue. *Harvard Business Review*, 2020.

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, 2019.

Manyu Jiang. The reason zoom calls drain your energy, Apr 2020. URL <https://www.bbc.com/worklife/article/20200421-why-zoom-video-chats-are-so-exhausting>.

Monica Landoni, Jerry Alan Fails, Theo Huibers, Natalia Kucirkova, Emiliana Murgia, and Maria Soledad Pera. “4 th kidrec workshop-” what does good look like?” from design, research, and practice to policy. In *Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts*, pages 103–110, 2020.

Ashlee Milton and Maria Soledad Pera. Evaluating information retrieval systems for kids. *arXiv preprint arXiv:2005.12992*, 2020.

Patricia A. Norberg, Daniel R. Horne, and David A. Horne. The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. *Journal of Consumer Affairs*, 41(1):100–126, 2007. ISSN 1745-6606.

Maria Soledad Pera, Jerry Alan Fails, Mirko Gelsomini, and Franca Garzotto. Building community: Report on kidrec workshop on children and recommender systems at recsys 2017. In *ACM SIGIR Forum*, volume 52, pages 153–161. ACM New York, NY, USA, 2018.