Children and Search Tools: Evaluation Remains Unclear

Brody Downs, Tyler French
brodydowns@u.boisestate.edu
tylerfrench@u.boisestate.edu
Computer Science – Boise State University
Boise, Idaho

Katherine Landau Wright
katherinewright@boisestate.edu
Literacy, Language, and Culture – Boise State University
Boise, Idaho

Maria Soledad Pera
solepera@boisestate.edu
People & Information Research Team – Computer Science – Boise State University
Boise, Idaho

Casey Kennington, Jerry Alan Fails
caseykennington@boisestate.edu
jerryfails@boisestate.edu
Computer Science – Boise State University
Boise, Idaho

ABSTRACT

As children search the internet for materials, they often turn to search engines that, unfortunately, offer children little support as they formulate queries to initiate the search process or examine resources for relevance. While some solutions have been proposed to address this, inherent to this issue is the need to evaluate the effectiveness of these solutions. We posit that the evaluation of the diverse aspects involved in the search process – from query suggestion generation to resource retrieval – requires a complex, multi-faceted approach that draws on evaluation methods utilized in human-computer interaction, information retrieval, natural language processing, education, and psychology.

KidRec '19, June 15, 2019, Boise, ID

CCS CONCEPTS

• **Social and professional topics** → **Children**; • **Information systems** → Web searching and information discovery; Personalization; • **Human-centered computing** → Participatory design.

KEYWORDS

children, web search, participatory design, evaluation, interdisciplinary

ACM Reference Format:
Brody Downs, Tyler French, Katherine Landau Wright, Maria Soledad Pera, and Casey Kennington, Jerry Alan Fails. 2019. Children and Search Tools: Evaluation Remains Unclear. In *KidRec '19: Workshop in International and Interdisciplinary Perspectives on Children & Recommender and Information Retrieval Systems, Co-located with ACM IDC, June 15, 2019, Boise, ID.* ACM, New York, NY, USA, 4 pages.

INTRODUCTION

Children are increasingly turning to the internet to find information for school and entertainment. They often use popular search engines to find the materials they need; however, these search tools are primarily designed for adults, not children. Consequently, they do not always support child-specific needs, particularly surrounding formulating queries or facilitating the understanding and analysis of the retrieved resources. Both of these types of assistance are important particularly for children, and some solutions have been suggested to address these issues [1, 3, 6, 8, 10].

Understanding what children need when it comes to successfully completing a search task rests in the intersection of research from human-computer interaction (**HCI**) which has focused on interfaces and user interaction, natural language processing (**NLP**) which has focused on uses of language, and information retrieval (**IR**) which has focused on search, retrieval, and ranking algorithms. Surrounding each facet of the search experience for children in this setting is **education** and **psychology** (particularly child development) to guide HCI, NLP, and IR research to help the children as they learn about the tools and content that the tools provide so the tools help children in their search goals.

Due of the nature of this multi-disciplinary setting, in this paper we propose approaching the evaluation of child search tools with a multi-pronged approach that includes elements from these research domains. (See Table 1 for an overview.) A high-level sketch is that evaluation requires regular formative assessments with end-users: children. We accomplish this via regular participatory design sessions with children. In the following sections we identify challenges in the field and conclude with our proposed approaches for evaluation.

Table 1: Measures to evaluate search tools.

- **Query Length.** Indicator of complexity of query and difficulty of user on expressing information needs
- **Click Analysis.** Rank distribution of resources retrieved; position click bias
- **Session Length.** Indicator of the average number of (new) queries employed to locate relevant resources
- **Navigation Style.** Identify (validate) strategy adopted by 6-11 year olds to solve search tasks
- **Session Duration.** Indicator of average time spent in locating relevant resources
- **Query Purpose.** Indicator of distribution of query purpose, i.e., navigational versus informational
- **Spelling.** Identify children's difficulty with spelling
- **Question Queries.** Indicator of percentage of queries posed using natural language
- **Search Suggestions.** Indicator of the degree to which targeted audience selects search suggestions
- **Readability.** Average readability level of snippets and clicked results
- **Task Completion.** Average time needed to accomplish a search task
- **Errors.** Average number of errors encountered; or reformulations needed to accomplish task
- **Satisfaction.** Average subjective satisfaction in finding results

CURRENT STATE OF EVALUATION FOR CHILD SEARCH TOOLS

Benchmarks

Researchers have widely documented the challenges that children face, but there are gaps in understanding and the technology research that address these challenges. In particular, a scarcity of evaluation tools hinders progress. To our knowledge, there are no common benchmarks or datasets for evaluating children's search tools regardless of the field conducting the evaluation. For example, NIST's Text Retrieval Conference archives datasets that researchers can use to evaluate search systems for average users; however, no equivalent datasets exist for evaluating children's search tools. As a result, researchers cannot replicate empirical studies, so there are few unbiased assessments comparing systems. These isolated evaluations constrain the research process and make it difficult for new systems to improve from lessons learned using established and replicable frameworks. Moreover, these evaluations are generally performed offline; i.e., the focus is on search results and less on the user experience with the search tools.

User Studies

User studies are essential to (1) gather feedback from children and teachers about issues affecting children's Web search experiences and (2) measure the degree of satisfaction children have when conducting online searches using the proposed components over existing search engines. Unfortunately, researchers have generally based existing empirical studies on child search to a limited number of participants [9]. Other studies use the existing America Online (AOL) query log as a basis [5], but queries in this log are outdated and, as reported in [7], there is a chance that researchers utilize queries mistakenly identified as child queries.

POSITION: PARTICIPATORY DESIGN WITH CHILDREN AND EDUCATORS

We propose an approach for effective benchmarks and user studies of child search by leveraging qualitative analysis to gather children's feedback during their search experience: researchers should take notes on child interactions with modules, as previously done in similar assessment studies [4]. Moreover, search tools should be evaluated using quantitative measures including (but not limited to) the measures in Table 1. In isolation, each of the measures capture directives for evaluation from IR and NLP. Coupled with the participatory design, these measures work with HCI to build a holistic understanding of how well search tools are working for children.

To incorporate education and child development into the evaluation strategy, it is important to consider the broader setting in which children are using the search tools. Involving teachers and librarians who often aid emergent searchers in their quest for information online, can help gain

insights on child satisfaction with search tools. Furthermore, educational approaches to aspects of the above measures can also guide the approach taken to improve the tools. For example, NLP has traditionally approached spelling correction by simple lookup dictionaries and employ edit distances to determine the list of suggested target words. This is not sufficient for children who are still in the process of learning how to spell because children make spelling mistakes that are specific to their developmental stages [2]. These stages are well documented in education research and should guide child-specific spell checkers.

CONCLUSION

Adult-oriented search tools are not suitable for children, and neither are traditional evaluation approaches in fields that isolate individual components of search tools (e.g., spell checkers, query formulation assistance, or ranking of retrieved resources). The best way to determine if a search tool is suitable for children is to involve children in the design and evaluation processes and by using established benchmarks guided by educational research.

REFERENCES

[1] 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, 5 (2017), 524–545.

[2] Donald R Bear, Marcia Invernizzi, Shane Templeton, Francine R Johnston, et al. 2012. *Words their way: Word study for phonics, vocabulary, and spelling instruction.* Pearson/Allyn and Bacon.

[3] Kevyn Collins-Thompson, Paul N Bennett, Ryen W White, Sebastian De La Chica, and David Sontag. 2011. Personalizing web search results by reading level. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 403–412.

[4] Allison Druin, Elizabeth Foss, Hilary Hutchinson, Evan Golub, and Leshell Hatley. 2010. Children's roles using keyword search interfaces at home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 413–422.

[5] Sergio Duarte Torres and Ingmar Weber. 2011. What and how children search on the web. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 393–402.

[6] Jerry Alan Fails, Oghenemaro Anuyah, Casey Kennington, Katherine Landau Wright, William Bigirimana, and Maria Soledad Pera. 2019. Query Formulation Assistance for Kids: What is Available, When to Help & What Kids Want. In *Proceedings of the 18th ACM Conference on Interaction Design and Children*. ACM, To appear.

[7] Tatiana Gossen. 2016. *Search engines for children: search user interfaces and information-seeking behaviour.* Springer.

[8] Hilary Browne Hutchinson. 2003. Children's Interface Design for Hierarchical Search and Browse. *SIGCAPH Comput. Phys. Handicap.* 75 (Jan. 2003), 11–12. https://doi.org/10.1145/976261.976265

[9] Yvonne Kammerer and Maja Bohnacker. 2012. Children's web search with Google: the effectiveness of natural language queries. In *proceedings of the 11th international conference on interaction design and children*. ACM, 184–187.

[10] Ion Madrazo Azpiazu, Nevena Dragovic, Oghenemaro Anuyah, and Maria Soledad Pera. 2018. Looking for the Movie Seven or Sven from the Movie Frozen?: A Multi-perspective Strategy for Recommending Queries for Children. In *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval*. ACM, 92–101.