

# Preface to the Proceedings of the Third International Workshop on Investigating Learning during Web Search (IWILDS'22)

Anett Hoppe<sup>3</sup>, Ran Yu<sup>4</sup> and Jiqun Liu<sup>5</sup>

<sup>3</sup>TIB – Leibniz Information Centre for Science and Technology & L3S Research Centre, Leibniz University Hannover, Germany

<sup>4</sup>Data Science & Intelligent Systems Group (DSIS), University of Bonn, Bonn, Germany

<sup>5</sup>The University of Oklahoma, Rono-Hills, Norman, OK, United States

## 1. Workshop summary

The Third International Workshop on Investigating Learning During Web Search (IWILDS 2022) was held in conjunction with the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2021), July 11-15, 2022 in Madrid, Spain. The program featured one keynote talk and the presentation of five accepted submissions (see Section 2 for more information on contents).

Web search is one of the most ubiquitous online activities and often used as a starting point to learn, i. e., to acquire or extend one's knowledge about certain topics or procedures. When learning by searching the Web, individuals are confronted with an unprecedented amount of information in various forms and varying quality. In consequence, successful learning on the Web is influenced by a range of individual and external factors [1]. It requires high degrees of self-regulation from the individual; and should be supported by the adequate design of search interfaces, recommendation engines, and training tools.


The question how to best support human learning on the Web creates a highly interdisciplinary research area at the intersection of information & multimedia retrieval, human-computer interaction, psychology, and educational sciences: *Search as Learning* (SAL) research examines the relationships between querying, navigation, media consumption behavior, and the learning outcomes during Web search, how they can be measured, predicted [2], and supported [3]. Leveraging the knowledge about learning processes and outcomes during search, SAL researchers also seek to incorporate learning factors into search evaluation and system design for supporting complex learning tasks [4, 5].


---

*Joint Proceedings of 10th International Workshop on News Recommendation and Analytics (INRA'22) and the Third International Workshop on Investigating Learning During Web Search (IWILDS'22) co-located with the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'22), July 15, 2022, Madrid, Spain*

✉ anett.hoppe@tib.eu (A. Hoppe); ran.yu@uni-bonn.de (R. Yu); jiqunliu@ou.edu (J. Liu)

🆔 0000-0002-1452-9509 (A. Hoppe); 0000-0002-1619-3164 (R. Yu); 0000-0003-3643-2182 (J. Liu)

 © 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

Building on the growing attention in SAL research, IWILDS covers multiple central research topics in SIGIR community (e.g. interpretation of user behavior [6] and user modeling [7], task understanding [8], search education [9], learning-centric re-ranking [10], datasets, measurement and evaluation [11, 12]) and provides an interdisciplinary forum in a half-day workshop that includes keynotes, paper presentations, and discussion.

Topics of interest include but are not limited to:

- **Understanding and measuring learning during Web search:** This includes works on the role of personal characteristics, interests, attitudes, and information literacy in Web-based learning; characterizing learning tasks; measuring the impacts of human cognitive biases and algorithmic biases on learning; methods of data collection and analysis (including crowd-sourcing, lab experiments, log analysis, multimodal data analysis); modeling, recognizing, measuring, and predicting learning processes; and the role of Web-based learning and search in formal and semi-formal scenarios.
- **Supporting learning during Web search:** e.g., interventions, tools, and user interfaces to foster effective SAL; information/ multimedia retrieval and learning to rank for SAL; learning analytics and educational data mining in search-based learning; fusion and summarization techniques for aggregating learning resources; evaluation and benchmarking of SAL systems; and personalization of retrieval and ranking in SAL.
- **Learning-oriented evaluation of Web search:** e.g., offline evaluation for learning-related ad hoc retrieval; designing and meta-evaluating evaluation metrics for capturing learning progresses in session search; evaluating learning activities in conversational IR; standardizing and reusing user study materials for replicable learning-oriented IR evaluation.

IWILDS 2022 received six submissions. Each paper has been reviewed by at least three reviewers in a single-blind reviewing process. Based on the reviews, five papers have been accepted for presentations. IWILDS 2022 took place with a hybrid setting with two of the workshop chairs onsite, and one chair attending virtually. The workshop has attracted 10-15 attendance onsite, and 20-25 attendance online.

The keynote speech *Understanding the “Pathway” Towards a Searcher’s Learning Objective* was given online by Jaime Arguello. He is an Associate Professor at the School of Information and Library Science at the University of North Carolina (UNC) at Chapel Hill. Jaime received his Ph.D. from the Language Technologies Institute at Carnegie Mellon University in 2011. Since then, his research has focused on a wide range of areas, including aggregated search, voice query reformulation, understanding search behaviors during complex tasks, developing search assistance tools to support complex tasks, and understanding the effects of specific cognitive abilities on search behaviors and outcomes. He has received Best Paper Awards at SIGIR 2009, ECIR 2011, IiX 2014, and ECIR 2017. His current research is supported by two NSF grants. Since 2015, Jaime has chaired the SIGIR Student Travel Grants Program and he recently served as PC co-chair for CHIIR 2022 and is an Associate Editor for TOIS.

Among the five paper talks in the workshop, two authors presented their work online: Kelsey Urgo presented *Capturing Self-Regulated Learning During Search*, and Markus Rokicki presented *Learning to Rank for Knowledge Gain*. Three other paper authors presented their work onsite:

Ratan Sebastian presented *Grade Level Filtering for Learning Object Search using Entity Linking*, Monica Landoni presented *Let's Learn from Children: Scaffolding to Enable Search as Learning in the Educational Environment*, and Abdolali Faraji presented *Goal-Driven Lifelong Learning through Personalized Search and Recommendation Services*. For details on each of the articles' contents, please refer to Section 2.

The talks presented in IWILDS'22 featured multidisciplinary research on topics including information retrieval, human computer interaction, semantic web and online education. Both theoretical and practical topics were discussed, fruitful discussion was triggered between the attending participants, both online and offline. We believe that the diverse domain expertise among IWILDS'22 attendance aligns well with the search as learning community. We believe the workshop served as a forum for interdisciplinary knowledge exchange and hope it enables further collaborations.

## 2. Accepted submissions

Information of the accepted papers are listed in this Section.

### **Capturing Self-Regulated Learning During Search**

**Kelsey Urgo, and Jaime Arguello.** Researchers in the learning sciences have demonstrated the benefits of effective self-regulated learning (SRL) in improving learning outcomes. The search-as-learning community aims to improve learning outcomes during search, but offers limited research exploring the impact of SRL on learning during search. Current limited research in search-as-learning explores only perceptions of SRL processes after the search process. Results from such analyses are limited in that SRL is a dynamic, active process and participant perceptions of SRL can be unreliable. In this paper, the authors propose the implementation of an SRL coding framework to capture SRL processes as they unfold throughout a search session. Additionally, they offer several implications for future work using the proposed methodology.

### **Learning to Rank for Knowledge Gain**

**Markus Rokicki, Ran Yu, and Daniel Hienert.** Web search has often been used as a starting point to learn. Search as Learning (SAL) research aims at supporting learning activities through techniques such as user interface optimization, retrieval, and ranking. In this work, the authors investigate the possibility of re-ranking search engine results towards learning to improve the overall knowledge gain of the learner. They make two contributions: (1) proposing a framework for re-ranking search results by attributing the overall knowledge gain to viewed documents in the session. (2) Applying this framework to a SAL evaluation dataset. They show that the ranking can be significantly improved with respect to knowledge gain by using ranking and content features.

## **Grade Level Filtering for Learning Object Search using Entity Linking**

**Ratan Sebastian, Ralph Ewerth, and Anett Hoppe** More and more Learning Objects are available online. To search and find suitable materials is, however, often a challenge as a lot of them lack sufficient metadata information concerning their format, content and appropriate application areas. One key piece of information about a Learning Object which is often missing is the targeted age bracket or grade level. This work studies the automatic content-based assignment of a resource's grade level. For this purpose, the authors (a) collected a dataset of physics Learning Objects, (b) explored a set of text-based features for their automatic analysis (derived from both dense vector representations and entity linking methods) and (c) trained a machine learning model with different subsets of these features to predict a resource's target grade level.

## **Let's Learn from Children: Scaffolding to Enable Search as Learning in the Educational Environment**

**Maria Soledad Pera, Monica Landoni, Emiliana Murgia, and Theo Huibers** In this manuscript, the authors argue for the need to further look at search as learning (SAL) with children as the primary stakeholders. Inspired by how children learn and considering the classroom (regardless of the teaching modality) as a natural educational ecosystem, we posit that scaffolding is the tie that can simultaneously allow for learning to search while searching for learning. The main contribution of this work is a list of open challenges focused on the primary school classroom for the IR community to consider when setting up to explore and make progress on SAL research with and for children and beyond.

## **Goal-Driven Lifelong Learning through Personalized Search and Recommendation Services**

**Abdolali Faraji, Mohammadreza Tavakoli and Gábor Kismihók** The need to keep your skills up to date is becoming more and more essential in the current, permanently changing educational world. At the same time, the number of published educational content on the web is continuously increasing, while the lack of metadata and proper quality control of these educational content is becoming an important issue for the search engine providers and educational recommender systems. This status quo is highly problematic for learners on the one hand when it comes to finding the most suitable educational material for their desired skills. On the other hand, this has also made the maintenance of curricula and learning pathways a frustrating job for content authors. In this research, the authors propose a novel Human-AI based recommender system, which combines a learning dashboard, and an open learning content/curriculum curation dashboard into one unified system to tackle the problem of individual learning path creation and maintenance both for content authors and lifelong learners.

### 3. Organising Committee

**Anett Hoppe** (Ph.D. Computer Science) is a postdoctoral researcher at Leibniz Information Centre for Science and Technology (TIB) and the L3S research Centre, both in Hannover, Germany. Her research focuses on the support of human learning with current technologies, with a focus on Information Retrieval, Video Analysis, User Modeling, and Semantic Web. She is involved in a number of interdisciplinary projects, collaborating with researchers from educational sciences and psychology. She has been a member of numerous program committees and co-organizer of SALMM'19, IWILDS'20 & '21 workshops.

**Jiqun Liu** (Ph.D. Information Science) is currently an assistant professor of data science and adjunct assistant professor of psychology at the University of Oklahoma. He holds a Ph.D. in Information Science from Rutgers University. His research focuses on the intersection of human-computer interaction (HCI), interactive information seeking/retrieval (IS&R), and cognitive psychology and seeks to apply the knowledge learned about people interacting with information in user modeling, adaptive recommendation, IR system evaluation, and intelligent nudging.

**Ran Yu** (Ph.D. Computer Science) is a senior researcher in the Data Science & Intelligent Systems (DSIS) research group at the University of Bonn. Her research interests are in Information Retrieval, User Modeling, Knowledge Graphs, and their application to Web data analytics problems, specifically in learning scenarios. Her work has been published in major conferences and journals; she is a member of numerous program committees such as SIGIR, WSDM, TheWebConf and CIKM, and has organized several academic events.

### 4. Programme committee

- Ralph Ewerth, L3S Research Center, Leibniz Universität Hannover, Germany
- Catherine Smith, Kent State University, USA
- Nilavra Bhattacharya, The University of Texas at Austin, USA
- Sherzod Hakimov, TIB – Leibniz Information Centre for Science and Technology, Germany
- Yvonne Kammerer, Stuttgart Media University, Germany
- Kevyn Collins-Thompson, University of Michigan, USA
- Bernardo Pereira Nunes, Australian National University, Australia
- Yuan Li, The University of North Carolina at Chapel Hill, USA
- Xiaolong Jin, Institute of Computing Technology, Chinese Academy of Sciences, China
- Sihang Qiu, Delft University of Technology, Netherlands
- Chang Liu, Peking University, China
- Gábor Kismihók, Leibniz Information Centre for Science and Technology, Germany

We thank our program committee members for producing outstanding reviews and all the attendees for the feedback and good discussions.

## References

- [1] J. von Hoyer, A. Hoppe, Y. Kammerer, C. Otto, G. Pardi, M. Rokicki, R. Yu, S. Dietze, R. Ewerth, P. Holtz, The search as learning spaceship: Toward a comprehensive model of psychological and technological facets of search as learning, *Frontiers in Psychology* 13 (2022). URL: <https://www.frontiersin.org/article/10.3389/fpsyg.2022.827748>. doi:10.3389/fpsyg.2022.827748.
- [2] C. Otto, R. Yu, G. Pardi, J. von Hoyer, M. Rokicki, A. Hoppe, P. Holtz, Y. Kammerer, S. Dietze, R. Ewerth, Predicting knowledge gain during web search based on multimedia resource consumption, in: I. Roll, D. S. McNamara, S. A. Sosnovsky, R. Luckin, V. Dimitrova (Eds.), *Artificial Intelligence in Education - 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14-18, 2021, Proceedings, Part I*, volume 12748 of *Lecture Notes in Computer Science*, Springer, 2021, pp. 318–330. URL: [https://doi.org/10.1007/978-3-030-78292-4\\_26](https://doi.org/10.1007/978-3-030-78292-4_26). doi:10.1007/978-3-030-78292-4\_26.
- [3] S. Y. Rieh, K. Collins-Thompson, P. Hansen, H. Lee, Towards searching as a learning process: A review of current perspectives and future directions, *J. Inf. Sci.* 42 (2016) 19–34. URL: <https://doi.org/10.1177/0165551515615841>. doi:10.1177/0165551515615841.
- [4] J. Liu, Y. J. Jung, Interest development, knowledge learning, and interactive IR: toward a state-based approach to search as learning, in: F. Scholer, P. Thomas, D. Elswailer, H. Joho, N. Kando, C. Smith (Eds.), *CHIIR '21: ACM SIGIR Conference on Human Information Interaction and Retrieval*, Canberra, ACT, Australia, March 14-19, 2021, ACM, 2021, pp. 239–248. URL: <https://doi.org/10.1145/3406522.3446015>. doi:10.1145/3406522.3446015.
- [5] C. L. Smith, S. Y. Rieh, Knowledge-context in search systems: Toward information-literate actions, in: L. Azzopardi, M. Halvey, I. Ruthven, H. Joho, V. Murdock, P. Qvarfordt (Eds.), *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval, CHIIR 2019, Glasgow, Scotland, UK, March 10-14, 2019*, ACM, 2019, pp. 55–62. URL: <https://doi.org/10.1145/3295750.3298940>. doi:10.1145/3295750.3298940.
- [6] M. Merkt, A. Hoppe, G. Bruns, R. Ewerth, M. Huff, Pushing the button: Why do learners pause online videos?, *Comput. Educ.* 176 (2022) 104355. URL: <https://doi.org/10.1016/j.compedu.2021.104355>. doi:10.1016/j.compedu.2021.104355.
- [7] K. Radinsky, K. M. Svore, S. T. Dumais, J. Teevan, A. Bocharov, E. Horvitz, Modeling and predicting behavioral dynamics on the web, in: A. Mille, F. Gandon, J. Misselis, M. Rabinovich, S. Staab (Eds.), *Proceedings of the 21st World Wide Web Conference 2012, WWW 2012, Lyon, France, April 16-20, 2012*, ACM, 2012, pp. 599–608. URL: <https://doi.org/10.1145/2187836.2187918>. doi:10.1145/2187836.2187918.
- [8] J. Liu, Deconstructing search tasks in interactive information retrieval: A systematic review of task dimensions and predictors, *Inf. Process. Manag.* 58 (2021) 102522. URL: <https://doi.org/10.1016/j.ipm.2021.102522>. doi:10.1016/j.ipm.2021.102522.
- [9] J. Liu, Y. Wang, S. Mandal, C. Shah, Exploring the immediate and short-term effects of peer advice and cognitive authority on web search behavior, *Inf. Process. Manag.* 56 (2019)

1010–1025. URL: <https://doi.org/10.1016/j.ipm.2019.02.011>. doi:10.1016/j.ipm.2019.02.011.

- [10] R. Syed, K. Collins-Thompson, Retrieval algorithms optimized for human learning, in: N. Kando, T. Sakai, H. Joho, H. Li, A. P. de Vries, R. W. White (Eds.), Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Shinjuku, Tokyo, Japan, August 7-11, 2017, ACM, 2017, pp. 555–564. URL: <https://doi.org/10.1145/3077136.3080835>. doi:10.1145/3077136.3080835.
- [11] K. Urgo, J. Arguello, Learning assessments in search-as-learning: A survey of prior work and opportunities for future research, *Inf. Process. Manag.* 59 (2022) 102821. URL: <https://doi.org/10.1016/j.ipm.2021.102821>. doi:10.1016/j.ipm.2021.102821.
- [12] C. Otto, M. Rokicki, G. Pardi, W. Gritz, D. Hienert, R. Yu, J. von Hoyer, A. Hoppe, S. Dietze, P. Holtz, Y. Kammerer, R. Ewerth, Sal-lightning dataset: Search and eye gaze behavior, resource interactions and knowledge gain during web search, *CoRR* abs/2201.02339 (2022). URL: <https://arxiv.org/abs/2201.02339>. arXiv:2201.02339.