

Introduction to the Special Issue on Search as Learning

Carsten Eickhoff¹, Jacek Gwizdka², Claudia Hauff³, and Jiyin He⁴

¹ ETH Zurich, Switzerland

² University of Texas at Austin, USA

³ Delft University of Technology, The Netherlands

⁴ Centrum Wiskunde & Informatica (CWI), The Netherlands

The Internet connects vast knowledge resources that are assumed to contain the necessary information to answer most general questions [11, 17]. However, even given this seemingly inexhaustible well of information, the act of learning from it consists of more than just looking-up and memorizing facts. Constructivist theoreticians including Piaget and Vygotsky argue that the learning process necessarily depends on the context of existing knowledge upon which the newly encountered factoids are built. The stronger the contextualization of new knowledge, the more effortless and effective learning is assumed to be.

Jean Piaget first proposed the theory of “Cognitive Development” that considers knowledge to be an actively constructed complex system of experience, stage of cognitive development, cultural background and personal history [16]. In other words, knowledge is derived from personal experience and ideas rather than an aggregation of loose facts and formulas. Building on Piaget’s theories, Langley and Johnson [7, 13] study order effects in incremental learning. The authors find that the order in which material is consumed has a significant influence on the overall learning rate and absolute knowledge retention. Kuhlthau *et al.* [12] discuss the importance of mediators who enable learners to go beyond the current limits of their understanding.

Despite the wide acceptance and demonstrated success of constructivist methods in pedagogics, common retrieval models do not explicitly manifest any notion of *contextual learning*. Document relevance is largely judged in isolation and list-wide ranking considerations rarely go beyond diversification efforts. Consequently, state-of-the-art search engines cannot be considered ideal learning environments.

Over the past decade, the study of users’ domain-specific knowledge has led to a wide number of innovations in user behavior understanding. White *et al.* [20] show that, within their area of expertise, domain experts search differently from non-experts. They are found to use a more diverse vocabulary of query terms and generally demonstrate a better understanding of the desired results to be retrieved, resulting in improved query formulation and result inspection performance as compared to laypeople.

Such domain expertise is by no means a static notion and a number of studies have noted considerable long-term changes [14, 21] as well as more gradual within-session developments [1, 2]. This is assumed to occur as a consequence

of repeatedly being exposed to in-domain information, *e.g.*, during Web search, and therefore having acquired increased domain expertise.

While these works laid foundations towards knowledge acquisition in search, the dedicated academic interest in construing searching as a learning process, in recognizing the importance of learning as a search outcome and in providing explicit support for learning while searching is a rather recent development. The initial ideas were presented at recent workshops [3–5], in a special issue of *Journal of Information Science* [6] as well as during a Dagstuhl Seminar⁵. This special issue continues these themes and brings them to the information retrieval community by presenting six novel research efforts into expertise-based search strategies and support, collaborative learning paradigms and retrieval technology that bridges the effectiveness gap between expert and novice searchers.

Accepted Articles

Kodama *et al.* [10] investigate the relationship between middle school students' mental models of Google and their information seeking skills. Students tended to personify Google and showed a limited understanding of how search engines work. This research suggests to developers and interface designers to make the inner workings of the engine more transparent and has implications on digital literacy instruction. Lu and Hsiao [15] study the information seeking behavior of users in programming language forums. The authors also devise a personalized information seeking assistant that fosters learning through user behavior modeling and query refinement, showing significantly improved learning effectiveness. Karanam *et al.* [8] present a search result click prediction model incorporating the user's individual level of relevant domain knowledge. The authors further discuss and compare strategies for knowledge acquisition tailored to the existing state of knowledge, showing significant gains in knowledge acquisition when using expertise-adequate strategies.

Knight *et al.* [9] investigate the usefulness of orchestration, a framework explicitly manifesting background context, actors and activities, for collaborative learning and searching. The authors' empirical evaluation is set in a classroom environment and promises beneficial outcomes for use in educational institutions.

Syed and Collins-Thompson [19] study retrieval methods that are optimized for enhancing the user's learning efficiency by diversifying relevant search results in terms of vocabulary usage. A crowdsourced user study demonstrates significant learning gains when rankers include a high density of previously unfamiliar new terms. Soledad-Pera *et al.* [18] present an enhanced search environment You-Understood.Me (YUM) that aims to support learning of K-9 children through retrieving documents that not only match children information needs but also their reading abilities. The article takes a user-centered approach and makes a distinction between searching to learn and learning to search.

⁵ <http://www.dagstuhl.de/17092>

Conclusion

The six contributions to this special issue span a wide range of domain-specific studies, demonstrating a significant potential for studying search as an information seeking and learning activity. Each of them makes a strong case for modeling learning intents and activities in the context of information search and retrieval. In the future, further empowering learning efforts by means of technology will critically depend on our ability as a community to recognize and adequately represent instances of knowledge acquisition as they occur in real-life information systems. Especially the study of user models and personalization, not just on a topical, but also a domain-expertise level will be a key factor in this regard.

References

1. Carsten Eickhoff, Sebastian Dungs, and Vu Tran. An eye-tracking study of query reformulation. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 13–22. ACM, 2015.
2. Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan Dumais. Lessons from the journey: A query log analysis of within-session learning. In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 223–232. ACM, 2014.
3. Luanne Freund, Jacek Gwizdka, Preben Hansen, Noriko Kando, and Soo Young Rieh. From searching to learning. In Maristella Agosti, Norbert Fuhr, Elaine Toms, and Pertti Vakkari, editors, *Evaluation Methodologies in Information Retrieval*, volume 13441 of *Dagstuhl Reports*, pages 102–105.
4. Luanne Freund, Jiyin He, Jacek Gwizdka, Noriko Kando, Preben Hansen, and Soo Young Rieh. Searching as learning (SAL) workshop 2014. In *Proceedings of the 5th Information Interaction in Context Symposium, IiX '14*, pages 7–7. ACM.
5. Jacek Gwizdka, Preben Hansen, Claudia Hauff, Jiyin He, and Noriko Kando. Search as learning (SAL) workshop 2016. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '16*, pages 1249–1250. ACM.
6. Preben Hansen and Soo Young Rieh. Editorial recent advances on searching as learning: An introduction to the special issue. 42(1):3–6.
7. Elaine B Johnson. *Contextual teaching and learning: What it is and why it's here to stay*. Corwin Press, 2002.
8. Saraschandra Karanam, Guillermo Jorge-Botana, Ricardo Olmos, and Herre van Oostendorp. Cognitive modeling of domain knowledge effects on information search behavior. *Information Retrieval*, XX(XX), 2017.
9. Simon Knight, Bart Rienties, Karen Littleton, Dirk Tempelaar, Matthew Mitsui, and Chirag Shah. The orchestration of a collaborative information seeking learning task. *Information Retrieval*, XX(XX), 2017.
10. Christie Kodama, Beth St.Jean, Mega Subramaniam, and Natalie Greene Taylor. "there's a creepy guy on the other end at google!": Engaging middle school students in a drawing activity to elicit their mental models of google. *Information Retrieval*, XX(XX), 2017.
11. Jona Kräenbring, Tika Monzon Penza, Joanna Gutmann, Susanne Muehlich, Oliver Zolk, Leszek Wojnowski, Renke Maas, Stefan Engelhardt, and Antonio Sarikas. Accuracy and completeness of drug information in wikipedia: a comparison with standard textbooks of pharmacology. *PloS one*, 9(9):e106930, 2014.

12. Carol Collier Kuhlthau. Seeking meaning. *Norwood, NJ: Ablex*, 1993.
13. Pat Langley. Order effects in incremental learning. *Learning in humans and machines: Towards an interdisciplinary learning science. Pergamon*, 136:137, 1995.
14. Jingjing Liu, Nicholas J Belkin, Xiangmin Zhang, and Xiaojun Yuan. Examining users' knowledge change in the task completion process. *Information Processing & Management*, 49(5):1058–1074, 2013.
15. Yihan Lu and I-Han Hsiao. Personalized information seeking assistant (pisa): From programming information seeking to learning. *Information Retrieval*, XX(XX), 2017.
16. Jean Piaget. *Piaget's theory*. Springer, 1976.
17. Nicola J Reavley, Andrew J Mackinnon, Amy J Morgan, Mario Alvarez-Jimenez, Sarah E Hetrick, Eoin Killackey, Barnaby Nelson, Rosemary Purcell, Marie BH Yap, and Anthony F Jorm. Quality of information sources about mental disorders: a comparison of wikipedia with centrally controlled web and printed sources. *Psychological medicine*, 42(08):1753–1762, 2012.
18. Maria Soledad-Pera, Ion Madrazo, Nevena Dragovic, and Jerry Alan Fails. Online searching and learning yum and other search tools for children and teachers. *Information Retrieval*, XX(XX), 2017.
19. Rohail Syed and Kevyn Collins-Thompson. Optimizing search results for human learning goals. *Information Retrieval*, XX(XX), 2017.
20. Ryen W White, Susan T Dumais, and Jaime Teevan. Characterizing the influence of domain expertise on web search behavior. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pages 132–141. ACM, 2009.
21. Barbara M Wildemuth. The effects of domain knowledge on search tactic formulation. *Journal of the american society for information science and technology*, 55(3):246–258, 2004.