

# Tutorial on Task-Based Search and Assistance

CHIRAG SHAH, University of Washington

RYEN W. WHITE, Microsoft Research

## ACM Reference Format:

Chirag Shah and Ryen W. White. 2020. Tutorial on Task-Based Search and Assistance. In *SIGIR '20: ACM SIGIR, July 25–30, 2020, Xi'an, China*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## REFERENCES

- [1] Maristella Agosti, Norbert Fuhr, Elaine Toms, and Pertti Vakkari. 2014. Evaluation methodologies in information retrieval dagstuhl seminar 13441. In *ACM SIGIR Forum*, Vol. 48. ACM New York, NY, USA, 36–41.
- [2] David Allen. 2015. *Getting Things Done: The Art of Stress-free Productivity*. Penguin.
- [3] Leif Azzopardi, Paul Thomas, and Nick Craswell. 2018. Measuring the utility of search engine result pages: an information foraging based measure. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 605–614.
- [4] Krisztian Balog. 2015. Task-completion Engines: A Vision with a Plan.. In *SCST@ ECIR*. Citeseer.
- [5] Nicholas Belkin, Toine Bogers, Jaap Kamps, Diane Kelly, Marijn Koolen, and Emine Yilmaz. 2017. Second workshop on supporting complex search tasks. In *Proceedings of the ACM SIGIR CHIIR Conference on Conference Human Information Interaction and Retrieval*. 433–435.
- [6] Victoria Bellotti, Brinda Dalal, Nathaniel Good, Peter Flynn, Daniel G Bobrow, and Nicolas Ducheneaut. 2004. What a to-do: Studies of task management towards the design of a personal task list manager. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 735–742.
- [7] Paul N Bennett, Ryen W White, Wei Chu, Susan T Dumais, Peter Bailey, Fedor Borisyyuk, and Xiaoyuan Cui. 2012. Modeling the impact of short-and long-term behavior on search personalization. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*. 185–194.
- [8] Marilyn G Boltz, Cara Kupperman, and Jessica Dunne. 1998. The role of learning in remembered duration. *Memory & Cognition* 26, 5 (1998), 903–921.
- [9] Pia Borlund. 2000. Experimental components for the evaluation of interactive information retrieval systems. *Journal of documentation* 56, 1 (2000), 71–90.
- [10] Paweł Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Inigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gašić. 2018. Multiwoz-a large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling. *arXiv preprint arXiv:1810.00278* (2018).
- [11] Katriina Byström and Preben Hansen. 2005. Conceptual framework for tasks in information studies. *Journal of the American Society for Information Science and Technology* 56, 10 (2005), 1050–1061.
- [12] Katriina Byström and Kalervo Järvelin. 1995. Task complexity affects information seeking and use. *Information Processing & Management* 31, 2 (1995), 191–213.
- [13] Robert Capra, Jaime Arguello, Heather O'Brien, Yuan Li, and Bogeum Choi. 2018. The effects of manipulating task determinability on search behaviors and outcomes. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*. 445–454.
- [14] Stuart K Card and Thomas P Moran. 1983. Newell, The psychology of human-computer interaction.
- [15] Eric L Charnov et al. 1976. Optimal foraging, the marginal value theorem. (1976).
- [16] Charles LA Clarke, Maheedhar Kolla, Gordon V Cormack, Olga Vechtomova, Azin Ashkan, Stefan Büttcher, and Ian MacKinnon. 2008. Novelty and diversity in information retrieval evaluation. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. 659–666.
- [17] William S Cooper. 1968. Expected search length: A single measure of retrieval effectiveness based on the weak ordering action of retrieval systems. *American documentation* 19, 1 (1968), 30–41.
- [18] Mary Czerwinski, Eric Horvitz, and Susan Wilhite. 2004. A diary study of task switching and interruptions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 175–182.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2020 Association for Computing Machinery.

Manuscript submitted to ACM

- [19] Henry A Feild, James Allan, and Rosie Jones. 2010. Predicting searcher frustration. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. 34–41.
- [20] Steve Fox, Kuldeep Karnawat, Mark Mydland, Susan Dumais, and Thomas White. 2005. Evaluating implicit measures to improve web search. *ACM Transactions on Information Systems (TOIS)* 23, 2 (2005), 147–168.
- [21] Qi Guo and Eugene Agichtein. 2012. Beyond dwell time: estimating document relevance from cursor movements and other post-click searcher behavior. In *Proceedings of the 21st international conference on World Wide Web*. 569–578.
- [22] Ahmed Hassan, Rosie Jones, and Kristina Lisa Klinkner. 2010. Beyond DCG: user behavior as a predictor of a successful search. In *Proceedings of the third ACM international conference on Web search and data mining*. 221–230.
- [23] Ahmed Hassan, Yang Song, and Li-wei He. 2011. A task level metric for measuring web search satisfaction and its application on improving relevance estimation. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. 125–134.
- [24] Ahmed Hassan and Ryen W White. 2013. Personalized models of search satisfaction. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*. 2009–2018.
- [25] Ahmed Hassan Awadallah, Cathal Gurrin, Mark Sanderson, and Ryen W White. 2019. Task Intelligence Workshop@ WSDM 2019. In *Proceedings of the ACM WSDM Conference on Web Search and Data Mining*. 848–849.
- [26] Jeff Huang, Ryen W White, Georg Buscher, and Kuansan Wang. 2012. Improving searcher models using mouse cursor activity. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*. 195–204.
- [27] Peter Ingwersen and Kalervo Järvelin. 2005. *The Turn: Integration of Information Seeking and Retrieval in Context*. Springer.
- [28] Richard David Jacques. 1996. *The nature of engagement and its role in hypermedia evaluation and design*. Ph.D. Dissertation. South Bank University.
- [29] Kalervo Järvelin, Susan L Price, Lois ML Delcambre, and Marianne Lykke Nielsen. 2008. Discounted cumulated gain based evaluation of multiple-query IR sessions. In *European Conference on Information Retrieval*. Springer, 4–15.
- [30] Jiepu Jiang, Ahmed Hassan Awadallah, Xiaolin Shi, and Ryen W White. 2015. Understanding and predicting graded search satisfaction. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. 57–66.
- [31] Daniel Kahneman and Amos Tversky. 1977. *Intuitive prediction: Biases and corrective procedures*. Technical Report. Decisions and Designs Inc Mclean Va.
- [32] Diane Kelly. 2009. Methods for evaluating interactive information retrieval systems with users. *Foundations and trends in Information Retrieval* 3, 1–2 (2009), 1–224.
- [33] Diane Kelly. 2015. When effort exceeds expectations: A theory of search task difficulty. In *ECIR Supporting Complex Search Task Workshop '15*.
- [34] Diane Kelly, Jaime Arguello, and Robert Capra. 2013. NSF workshop on task-based information search systems. In *ACM SIGIR Forum*, Vol. 47. 116–127.
- [35] Diane Kelly, Jaime Arguello, Ashlee Edwards, and Wan-ching Wu. 2015. Development and evaluation of search tasks for IIR experiments using a cognitive complexity framework. In *Proceedings of the ACM SIGIR ICTIR International Conference on the Theory of Information Retrieval*. ACM, 101–110.
- [36] Diane Kelly and Nicholas J Belkin. 2004. Display time as implicit feedback: understanding task effects. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*. 377–384.
- [37] Diane Kelly and Colleen Cool. 2002. The effects of topic familiarity on information search behavior. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*. 74–75.
- [38] Joyce Diane Kelly. 2004. *Understanding implicit feedback and document preference: A naturalistic user study*. Ph.D. Dissertation. Rutgers the State University of New Jersey-New Brunswick.
- [39] Youngho Kim, Ahmed Hassan, Ryen W White, and Imed Zitouni. 2014. Modeling dwell time to predict click-level satisfaction. In *Proceedings of the 7th ACM international conference on Web search and data mining*. 193–202.
- [40] Julia Kiseleva, Kyle Williams, Jiepu Jiang, Ahmed Hassan Awadallah, Aidan C Crook, Imed Zitouni, and Tasos Anastasakos. 2016. Understanding user satisfaction with intelligent assistants. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*. 121–130.
- [41] Alexander Kotov, Paul N Bennett, Ryen W White, Susan T Dumais, and Jaime Teevan. 2011. Modeling and analysis of cross-session search tasks. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. 5–14.
- [42] Mounia Lalmas, Heather O'Brien, and Elad Yom-Tov. 2014. Measuring user engagement. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 6, 4 (2014), 1–132.
- [43] Jane Li, Scott Huffman, and Akihito Tokuda. 2009. Good abandonment in mobile and PC internet search. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. 43–50.
- [44] Yuelin Li and Nicholas J Belkin. 2008. A faceted approach to conceptualizing tasks in information seeking. *Information Processing & Management* 44, 6 (2008), 1822–1837.
- [45] Jonathan Liono, Mohammad Saiedur Rahaman, Flora D Salim, Yongli Ren, Damiano Spina, Falk Scholer, Johanne R Trippas, Mark Sanderson, Paul N Bennett, and Ryen W White. 2020. Intelligent Task Recognition: Towards Enabling Productivity Assistance in Daily Life. In *Proceedings of the 2020 International Conference on Multimedia Retrieval*. 472–478.
- [46] Chia-Wei Liu, Ryan Lowe, Iulian V Serban, Michael Noseworthy, Laurent Charlin, and Joelle Pineau. 2016. How not to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. *arXiv preprint arXiv:1603.08023* (2016).

- [47] Jiqun Liu, Matthew Mitsui, Nicholas J Belkin, and Chirag Shah. 2019. Task, information seeking Intentions, and user behavior: Toward a multi-level understanding of web search. In *Proceedings of the ACM SIGIR CHIIR Conference on Human Information Interaction and Retrieval*. 123–132.
- [48] Jiqun Liu, Matthew Mitsui, Nicholas J. Belkin, and Chirag Shah. 2019. Task, information seeking intentions, and user behavior: Toward a multi-level understanding of web search. In *CHIIR 2019 - Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*. <https://doi.org/10.1145/3295750.3298922>
- [49] Shane J Lopez and Charles R Snyder. 2009. *The Oxford handbook of positive psychology*. Oxford University Press.
- [50] Rishabh Mehrotra, A Hassan Awadallah, Ahmed E Kholy, and Imed Zitouni. 2017. Hey Cortana! Exploring the use cases of a Desktop based Digital Assistant. In *SIGIR 1st International Workshop on Conversational Approaches to Information Retrieval (CAIR'17)*, Vol. 4.
- [51] Nina Mishra, Ryen W White, Samuel leong, and Eric Horvitz. 2014. Time-critical search. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. 747–756.
- [52] Matthew Mitsui, Jiqun Liu, and Chirag Shah. 2018. How much is too much?: Whole session vs. First query behaviors in task type prediction. In *41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018*. <https://doi.org/10.1145/3209978.3210105>
- [53] Matthew Mitsui, Jiqun Liu, and Chirag Shah. 2018. The paradox of personalization: Does task prediction require individualized models?. In *CHIIR 2018 - Proceedings of the 2018 Conference on Human Information Interaction and Retrieval*. <https://doi.org/10.1145/3176349.3176887>
- [54] Matthew Mitsui and Chirag Shah. 2018. The Broad View of Task Type Using Path Analysis. In *Proceedings of the 2018 ACM SIGIR International Conference on Theory of Information Retrieval*. ACM, 131–138.
- [55] Matthew Mitsui and Chirag Shah. 2018. The broad view of task type using path analysis. In *ICTIR 2018 - Proceedings of the 2018 ACM SIGIR International Conference on the Theory of Information Retrieval*. <https://doi.org/10.1145/3234944.3234951>
- [56] Matthew Mitsui and Chirag Shah. 2019. Bridging gaps: Predicting user and task characteristics from partial user information. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*. 415–424.
- [57] Matthew Mitsui and Chirag Shah. 2019. Bridging Gaps: Predicting User and Task Characteristics from Partial User Information. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*. ACM.
- [58] Stefano Mizzaro. 1997. Relevance: The whole history. *Journal of the American society for information science* 48, 9 (1997), 810–832.
- [59] Alistair Moffat, Peter Bailey, Falk Scholer, and Paul Thomas. 2017. Incorporating user expectations and behavior into the measurement of search effectiveness. *ACM Transactions on Information Systems (TOIS)* 35, 3 (2017), 1–38.
- [60] Karen Myers, Pauline Berry, Jim Blythe, Ken Conley, Melinda Gervasio, Deborah L McGuinness, David Morley, Avi Pfeffer, Martha Pollack, and Milind Tambe. 2007. An intelligent personal assistant for task and time management. *AI Magazine* 28, 2 (2007), 47–47.
- [61] Heather L O'Brien and Elaine G Toms. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology* 59, 6 (2008), 938–955.
- [62] Daan Odijk, Ryen W White, Ahmed Hassan Awadallah, and Susan T Dumais. 2015. Struggling and success in web search. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. 1551–1560.
- [63] Richard L Oliver. 2014. *Satisfaction: A behavioral perspective on the consumer: A behavioral perspective on the consumer*. Routledge.
- [64] Daniel M Russell, Diane Tang, Melanie Kellar, and Robin Jeffries. 2009. Task behaviors during web search: The difficulty of assigning labels. In *2009 42nd Hawaii International Conference on System Sciences*. IEEE, 1–5.
- [65] Miamaria Saastamoinen and Kalervo Järvelin. 2017. Search task features in work tasks of varying types and complexity. *Journal of the Association for Information Science and Technology* 68, 5 (2017), 1111–1123.
- [66] Reijo Savolainen. 2012. Expectancy-value beliefs and information needs as motivators for task-based information seeking. *Journal of Documentation* 68, 4 (2012), 492–511.
- [67] Ben Shneiderman. 2000. Creating creativity: user interfaces for supporting innovation. *ACM Transactions on Computer-Human Interaction (TOCHI)* 7, 1 (2000), 114–138.
- [68] Mark D Smucker and Charles LA Clarke. 2012. Time-based calibration of effectiveness measures. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*. 95–104.
- [69] Arthur R Taylor, Colleen Cool, Nicholas J Belkin, and William J Amadio. 2007. Relationships between categories of relevance criteria and stage in task completion. *Information Processing & Management* 43, 4 (2007), 1071–1084.
- [70] Kevin Thomas, Simon Handley, and Stephen Newstead. 2004. The effects of prior experience on estimating the duration of simple tasks. *Current Psychology of Cognition* 22, 2 (2004), 83–100.
- [71] Pertti Vakkari. 1999. Task complexity, problem structure and information actions: Integrating studies on information seeking and retrieval. *Information Processing & Management* 35, 6 (1999), 819–837.
- [72] Pertti Vakkari. 2001. A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of Documentation* 57, 1 (2001).
- [73] Yu Wang, Xiao Huang, and Ryen W White. 2013. Characterizing and supporting cross-device search tasks. In *Proceedings of the sixth ACM international conference on Web search and data mining*. 707–716.
- [74] Ryen W White. 2016. *Interactions with Search Systems*. Cambridge University Press.
- [75] Ryen W White. 2018. Skill discovery in virtual assistants. *Commun. ACM* 61, 11 (2018), 106–113.
- [76] Ryen W White, Paul N Bennett, and Susan T Dumais. 2010. Predicting short-term interests using activity-based search context. In *Proceedings of the 19th ACM international conference on Information and knowledge management*. 1009–1018.

- [77] Ryen W White, Susan T Dumais, and Jaime Teevan. 2009. Characterizing the influence of domain expertise on web search behavior. In *Proceedings of the second ACM international conference on web search and data mining*. 132–141.
- [78] Ryen W White, Adam Fourney, Allen Herring, Paul N Bennett, Nirupama Chandrasekaran, Robert Sim, Elnaz Nouri, and Mark J Encarnación. 2019. Multi-device digital assistance. *Commun. ACM* 62, 10 (2019), 28–31.
- [79] Ryen W White and Ahmed Hassan Awadallah. 2019. Task duration estimation. In *Proceedings of the ACM WSDM Conference on Web Search and Data Mining*. 636–644.
- [80] Ryen W White, Ahmed Hassan Awadallah, and Robert Sim. 2019. Task completion detection: A study in the context of intelligent systems. In *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*. 405–414.
- [81] Ryen W White, Ashish Kapoor, and Susan T Dumais. 2010. Modeling long-term search engine usage. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 28–39.
- [82] Barbara M Wildemuth, Luanne Freund, and Elaine G Toms. 2014. Studies of search task complexity or difficulty. *Science & Technology* 62, 9 (2014), 1676–1695.
- [83] Ya Xu and David Mease. 2009. Evaluating web search using task completion time. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. 676–677.
- [84] Dan Zakay and Richard A Block. 1996. The role of attention in time estimation processes. In *Advances in psychology*. Vol. 115. Elsevier, 143–164.