

Bibliography

- Ageev, M., Lagun, D., and Agichtein, E. (2013). Improving search result summaries by using searcher behavior data. In *Proceedings of the 35th ACM SIGIR*, pages 13–22.
- Agosto, D. E. (2002). Bounded rationality and satisficing in young people’s web-based decision making. *Journal of the Association for Information Science and Technology*, 53(1):16–27.
- Ali, H., Scholer, F., Thom, J. A., and Wu, M. (2009). User interaction with novel web search interfaces. In *Proceedings of the 21st OZCHI*, pages 301–304.
- Ashkan, A., Clarke, C. L., Agichtein, E., and Guo, Q. (2009). Classifying and characterizing query intent. In *Proceedings of the 31st ECIR*, pages 578–586.
- Athukorala, K., Oulasvirta, A., Glowacka, D., Vreeken, J., and Jacucci, G. (2014). Narrow or broad?: Estimating subjective specificity in exploratory search. In *Proceedings of the 23rd ACM CIKM*, pages 819–828.
- Azzopardi, L. (2009). Query side evaluation: An empirical analysis of effectiveness and effort. In *Proceedings of the 32nd ACM SIGIR*, pages 556–563.
- Azzopardi, L. (2011). The economics in interactive information retrieval. In *Proceedings of the 34th ACM SIGIR*, pages 15–24.
- Azzopardi, L., de Rijke, M., and Balog, K. (2007). Building simulated queries for known-item topics: An analysis using six European languages. In *Proceedings of the 30th ACM SIGIR*, pages 455–462.

- Azzopardi, L., Järvelin, K., Kamps, J., and Smucker, M. D. (2011). Report on the SIGIR 2010 workshop on the simulation of interaction. *SIGIR Forum*, 44(2):35–47.
- Azzopardi, L., Kelly, D., and Brennan, K. (2013). How query cost affects search behavior. In *Proceedings of the 36th ACM SIGIR*, pages 23–32.
- Azzopardi, L., Thomas, P., and Craswell, N. (2018). Measuring the utility of search engine result pages: An information foraging based measure. In *Proceedings of the 41st ACM SIGIR*, pages 605–614.
- Azzopardi, L. and Vinay, V. (2008). Retrievability: An evaluation measure for higher order information access tasks. In *Proceedings of the 17th ACM CIKM*, pages 561–570.
- Azzopardi, L. and Zuccon, G. (2015). An analysis of theories of search and search behavior. In *Proceedings of the 1st ACM ICTIR*, pages 81–90.
- Baillie, M., Azzopardi, L., and Crestani, F. (2006). Adaptive query-based sampling of distributed collections. In *Proceedings of the 13th SPIRE*, pages 316–328.
- Banks, J., Carson, J., and Nelson, B. (1996). *Discrete-event System Simulation*. Prentice-Hall international series in industrial and systems engineering. Prentice Hall.
- Baron, J., Beattie, J., and Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, 42(1):88–110.
- Bartlett, F. C. and Burt, C. (1933). Remembering: A study in experimental and social psychology. *British Journal of Educational Psychology*, 3(2):187–192.
- Baskaya, F., Keskustalo, H., and Järvelin, K. (2012). Time drives interaction: Simulating sessions in diverse searching environments. In *Proceedings of the 35th ACM SIGIR*, pages 105–114.
- Baskaya, F., Keskustalo, H., and Järvelin, K. (2013). Modeling behavioral factors in interactive information retrieval. In *Proceedings of the 22nd ACM CIKM*, pages 2297–2302.

- Bast, H. and Celikik, M. (2014). Efficient index-based snippet generation. *ACM Transactions on Information Systems*, 32(2):6:1–6:24.
- Bates, M. (1984). The fallacy of the perfect thirty-item online search. *RQ*, 24(1):pp. 43–50.
- Bates, M. J. (1989a). The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13(5):407–424.
- Bates, M. J. (1989b). Training and education for online. chapter Information Search Tactics, pages 96–105. Taylor Graham Publishing.
- Belkin, N. (1980). Anomalous states of knowledge as a basis for information retrieval. In *Canadian Journal of Information Science*, volume 5, pages 133–143.
- Benov, D. M. (2016). The manhattan project, the first electronic computer and the monte carlo method. *Monte Carlo Methods and Applications*, 22(1):73–79.
- Benson, P. G., Curley, S. P., and Smith, G. F. (1995). Belief assessment: An underdeveloped phase of probability elicitation. *Management Science*, 41(10):1639–1653.
- Berners-Lee, T., Dimitroyannis, D., Mallinckrodt, A. J., McKay, S., et al. (1994). World wide web. *Computers in Physics*, 8(3):298–299.
- Berryman, J. (2006). What defines “enough” information? how policy workers make judgments and decisions during information seeking: Preliminary results from an exploratory study. *Information Research: An International Electronic Journal*, 11(4):4.
- Boole, G. (1847). *The Mathematical Analysis of Logic*. Philosophical Library.
- Borlund, P. (2000). Evaluation of interactive information retrieval systems. Unpublished doctoral dissertation, Åbo Akademi University.
- Borlund, P. (2003). The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Information Research*, 8(5).

- Borlund, P. and Schneider, J. W. (2010). Reconsideration of the simulated work task situation: A context instrument for evaluation of information retrieval interaction. In *Proceedings of the 3rd IliX*, pages 155–164.
- Browne, G. J. and Pitts, M. G. (2004). Stopping rule use during information search in design problems. *Organizational Behavior and Human Decision Processes*, 95(2):208 – 224.
- Browne, G. J., Pitts, M. G., and Wetherbe, J. C. (2005). Stopping rule use during web-based search. In *Proceedings of the 38th HICSS*, pages 271b–271b.
- Brutlag, J. (2009). Speed matters for Google Web Search. <http://goo.gl/t7qGN8> (retrieved on March 14th, 2018).
- Busemeyer, J. R. (1982). Choice behavior in a sequential decision-making task. *Organizational Behavior and Human Performance*, 29(2):175 – 207.
- Busemeyer, J. R. and Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology*, 32(2):91–134.
- Callan, J., Allan, J., Clarke, C. L. A., Dumais, S., Evans, D. A., Sanderson, M., and Zhai, C. (2007). Meeting of the minds: An information retrieval research agenda. *SIGIR Forum*, 41(2):25–34.
- Card, S., Pirolli, P., Van Der Wege, M., Morrison, J., Reeder, R., Schraedley, P., and Boshart, J. (2001). Information scent as a driver of web behavior graphs: Results of a protocol analysis method for web usability. In *Proceedings of the 19th ACM CHI*, pages 498–505.
- Carr, N. (2008). Is google making us stupid? *Yearbook of the National Society for the Study of Education*, 107(2):89–94.
- Carterette, B. (2011). System effectiveness, user models, and user utility: A conceptual framework for investigation. In *Proceedings of the 34th ACM SIGIR*, pages 903–912.
- Carterette, B., Bah, A., and Zengin, M. (2015). Dynamic test collections for retrieval evaluation. In *Proceedings of the 5th ACM ICTIR*, pages 91–100.

- Carterette, B., Kanoulas, E., and Yilmaz, E. (2011). Simulating simple user behavior for system effectiveness evaluation. In *Proceedings of the 20th ACM CIKM*, pages 611–620.
- Chapelle, O., Metzler, D., Zhang, Y., and Grinspan, P. (2009). Expected reciprocal rank for graded relevance. In *Proceedings of the 18th ACM CIKM*, pages 621–630.
- Charles-Dominique, P. and Martin, R. (1972). *Behaviour and Ecology of Nocturnal Prosimians: Field Studies in Gabon and Madagascar*. Advances in ethology. P. Parey.
- Charnov, E. (1976). Optimal foraging, the Marginal Value Theorem. *Theoretical Population Biology*, 9(2):129–136.
- Chen, D., Chen, W., Wang, H., Chen, Z., and Yang, Q. (2012). Beyond ten blue links: Enabling user click modeling in federated web search. In *Proceedings of the 5th ACM WSDM*, pages 463–472.
- Chen, H. and Dumais, S. (2000). Bringing order to the web: Automatically categorizing search results. In *Proceedings of the 18th ACM CHI*, pages 145–152.
- Chen, M. C., Anderson, J. R., and Sohn, M. H. (2001). What can a mouse cursor tell us more?: Correlation of eye/mouse movements on web browsing. In *Proceedings of the 19th ACM CHI Extended Abstracts*, pages 281–282.
- Chen, P. P.-S. (1976). The entity-relationship model – toward a unified view of data. *ACM Transactions on Database Systems*, 1(1):9–36.
- Chi, E. H., Pirolli, P., Chen, K., and Pitkow, J. (2001). Using information scent to model user information needs and actions and the web. In *Proceedings of the 19th ACM CHI 2001*, pages 490–497.
- Chierichetti, F., Kumar, R., and Raghavan, P. (2011). Optimizing two-dimensional search results presentation. In *Proceedings of the 4th ACM WSDM*, pages 257–266.
- Chuklin, A., Markov, I., and de Rijke, M. (2015). *Click Models for Web Search*. Morgan & Claypool.

- Chuklin, A. and Serdyukov, P. (2012). Good abandonments in factoid queries. In *Proceedings of the 21st WWW*, pages 483–484.
- Clarke, C. L., Kolla, M., Cormack, G. V., Vechtomova, O., Ashkan, A., Büttcher, S., and MacKinnon, I. (2008). Novelty and diversity in information retrieval evaluation. In *Proceedings of the 31st ACM SIGIR*, pages 659–666.
- Clarke, C. L. A., Agichtein, E., Dumais, S., and White, R. W. (2007). The influence of caption features on clickthrough patterns in web search. In *Proceedings of the 30th ACM SIGIR*, pages 135–142.
- Cleverdon, C., Mills, J., and Keen, M. (1966). *Factors Determining the Performance of Indexing Systems*, volume 1:2 of *Factors Determining the Performance of Indexing Systems*.
- Cleverdon, C. W. (1991). The significance of the cranfield tests on index languages. In *Proceedings of the 14th ACM SIGIR*, pages 3–12.
- Collins-Thompson, K., Callan, J., Terra, E., and Clarke, C. L. (2004). The effect of document retrieval quality on factoid question answering performance. In *Proceedings of the 27th ACM SIGIR*, pages 574–575.
- Collins-Thompson, K., Hansen, P., and Hauff, C. (2017). Search as learning (dagstuhl seminar 17092). In *Dagstuhl Reports*, volume 7.
- Cooper, W. (1971). A definition of relevance for information retrieval. *Information Storage and Retrieval*, 7(1):19–37.
- Cooper, W. S. (1968). Expected search length: A single measure of retrieval effectiveness based on the weak ordering action of retrieval systems. *American Documentation*, 19(1):30–41.
- Cooper, W. S. (1973a). On selecting a measure of retrieval effectiveness. *Journal of the American Society for Information Science*, 24(2):87–100.

- Cooper, W. S. (1973b). On selecting a measure of retrieval effectiveness part ii. implementation of the philosophy. *Journal of the American Society for Information Science*, 24(6):413–424.
- Craswell, N., Zoeter, O., Taylor, M., and Ramsey, B. (2008). An experimental comparison of click position-bias models. In *Proceedings of the 1st ACM WSDM 2008*, pages 87–94.
- Crescenzi, A., Capra, R., and Arguello, J. (2013). Time pressure, user satisfaction and task difficulty. In *Proceedings of the 76th ASIS&T*.
- Crescenzi, A., Kelly, D., and Azzopardi, L. (2016). Impacts of time constraints and system delays on user experience. In *Proceedings of the 1st ACM CHIIR*, pages 141–150.
- Croft, B., Metzler, D., and Strohman, T. (2009). *Search Engines: Information Retrieval in Practice*. Addison-Wesley Publishing Company, USA.
- Cutrell, E. and Guan, Z. (2007). What are you looking for?: An eye-tracking study of information usage in web search. In *Proceedings of the 25th ACM CHI*, pages 407–416.
- Cutts, Q., Connor, R., Michaelson, G., and Donaldson, P. (2014). Code or (not code): Separating formal and natural language in cs education. In *Proceedings of the 9th WiPSCE*, pages 20–28.
- Das Sarma, A., Gollapudi, S., and Jeong, S. (2008). Bypass rates: Reducing query abandonment using negative inferences. In *Proceedings of the 14th ACM KDD*, pages 177–185.
- Dewey, M. (1891). Decimal classification and relative index for libraries, clippings, notes, etc. 240(41):407–593.
- Diriye, A., White, R., Buscher, G., and Dumais, S. (2012). Leaving so soon?: Understanding and predicting web search abandonment rationales. In *Proceedings of the 21st ACM CIKM*, pages 1025–1034.
- Dolamic, L. and Savoy, J. (2010). When stopword lists make the difference. *Journal of the Association for Information Science and Technology*, 61(1):200–203.

- Dostert, M. and Kelly, D. (2009). Users' stopping behaviors and estimates of recall. In *Proceedings of the 32nd ACM SIGIR*, pages 820–821.
- Dumais, S., Cutrell, E., and Chen, H. (2001). Optimizing search by showing results in context. In *Proceedings of the 19th ACM CHI*, pages 277–284.
- Edwards, A., Kelly, D., and Azzopardi, L. (2015). The impact of query interface design on stress, workload and performance. In *Proceedings of the 37th ECIR*, pages 691–702.
- Efthimiadis, E. N. (2000). Interactive query expansion: A user-based evaluation in a relevance feedback environment. *Journal of the American Society for Information Science*, 51(11):989–1003.
- Eliot, S. and Rose, J. (2009). *A Companion to the History of the Book*. Blackwell Companions to Literature and Culture. John Wiley & Sons.
- Farquhar, P. H. and Pratkanis, A. R. (1993). Decision structuring with phantom alternatives. *Management Science*, 39(10):1214–1226.
- Feild, H., Jones, R., Miller, R., Nayak, R., Churchill, E., and Velipasaoglu, E. (2010). Logging the search self-efficacy of amazon mechanical turkers. In *Proceedings of the CSE SIGIR Workshop*, pages 27–30.
- Fischhoff, B. (1977). Cost benefit analysis and the art of motorcycle maintenance. *Policy Sciences*, 8(2):177–202.
- Fischhoff, B., Slovic, P., and Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, 4(2):330.
- Fishwick, P. A. (1995). Computer simulation: The art and science of digital world construction. Technical report, University of Florida.
- Fox, C. (1992). Information retrieval. chapter Lexical Analysis and Stoplists, pages 102–130.

- Francis, W. and Kučera, H. (1979). *Manual of Information to Accompany A Standard Corpus of Present-day Edited American English, for Use with Digital Computers*. Brown University, Department of Linguistics.
- Francis, W. and Kučera, H. (1985). Frequency analysis of english usage: Lexicon and grammar. *Journal of English Linguistics*, 18(1):64–70.
- Fuhr, N. (2008). A probability ranking principle for interactive information retrieval. *Information Retrieval*, 11(3):251–265.
- Fuhr, N. and Lalmas, M. (2006). Advances in xml retrieval: The inex initiative. In *Proceedings of IWRIDL*, page 16.
- Gettys, C. F. and Fisher, S. (1979). Hypothesis plausibility and hypothesis generation. *Organizational Behavior and Human Performance*, 24(1):93–110.
- Gibb, J. A. (1958). Predation by tits and squirrels on the eucosmid *ernarmonia conicolana* (heyl.). *Journal of Animal Ecology*, 27(2):375–396.
- Gigerenzer, G. and Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In *Simple heuristics that make us smart*, pages 75–95. Oxford University Press.
- Green, R. (1984). Stopping rules for optimal foragers. *The American Naturalist*, 123(1):30–43.
- Guo, F., Liu, C., Kannan, A., Minka, T., Taylor, M., Wang, Y., and Faloutsos, C. (2009). Click chain model in web search. In *Proceedings of the 18th WWW*, pages 11–20.
- Hagen, M., Michel, M., and Stein, B. (2015). What was the query? generating queries for document sets with applications in cluster labeling. In *Natural Language Processing and Information Systems*, pages 124–133.
- Hagen, M., Michel, M., and Stein, B. (2016). Simulating ideal and average users. In *Proceedings of the 12th AIRS*, pages 138–154.
- Harman, D. (1993). Overview of the first trec conference. In *Proceedings of the 16th ACM SIGIR, SIGIR '93*, pages 36–47.

- Harman, D. (2010). Is the cranfield paradigm outdated? In *Proceedings of the 33rd ACM SIGIR*, page 1.
- Harper, D. J. and Kelly, D. (2006). Contextual relevance feedback. In *Proceedings of the 1st ACM IliX*, pages 129–137.
- Harrower, M. and Brewer, C. A. (2003). Colorbrewer.org: an online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37.
- Harvey, M. and Pointon, M. (2017). Searching on the go: the effects of fragmented attention on mobile web search tasks. In *Proceedings of the 40th ACM SIGIR*, pages 155–164.
- Hassan, A., Shi, X., Craswell, N., and Ramsey, B. (2013). Beyond clicks: Query reformulation as a predictor of search satisfaction. In *Proceedings of the 22nd CIKM*, pages 2019–2028.
- Hassan, A. and White, R. (2013). Personalized models of search satisfaction. In *Proceedings of the 22nd ACM CIKM*, pages 2009–2018.
- He, J., Duboue, P., and Nie, J.-Y. (2012). Bridging the gap between intrinsic and perceived relevance in snippet generation. In *Proceedings of COLING 2012*, pages 1129–1146.
- Hearst, M. (2009). *Search user interfaces*. Cambridge University Press.
- Hearst, M. A. (1995). Tilebars: Visualization of term distribution information in full text information access. In *Proceedings of the 13th ACM SIGCHI*, pages 59–66.
- Hearst, M. A. (1997). Texttiling: Segmenting text into multi-paragraph subtopic passages. *Comput. Linguist.*, 23(1):33–64.
- Heermann, D. W. (1990). *Computer-Simulation Methods*, pages 8–12.
- Heine, M. D. (1981). Simulation, and simulation experiments. In Spärck Jones, K., editor, *Information Retrieval Experiments*, pages 197–198. Butterworth-Heinemann.
- Hiemstra, D. (2009). *Information Retrieval Models*, pages 1–19.

- Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2):79–102.
- Huang, J., White, R. W., and Dumais, S. (2011). No clicks, no problem: Using cursor movements to understand and improve search. In *Proceedings of the 29th ACM CHI*, pages 1225–1234.
- Ingwersen, P. and Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context*. Springer Publishing Company, Incorporated.
- Iwasa, Y., Higashi, M., and Yamamura, N. (1981). Prey distribution as a factor determining the choice of optimal foraging strategy. *The American Naturalist*, 117(5):710–723.
- Iwata, M., Sakai, T., Yamamoto, T., Chen, Y., Liu, Y., Wen, J.-R., and Nishio, S. (2012). Aspects: Tile-based visualization of diversified web search results. In *Proceedings of the 35th ACM SIGIR*, pages 85–94.
- Jahoda, G. (1961). Electronic searching. volume 4 of *The state of the library art*, pages 139–320. Graduate School of Library Service, Rutgers University.
- Janetos, A. C. and Cole, B. J. (1981). Imperfectly optimal animals. *Behavioral Ecology and Sociobiology*, 9(3):203–209.
- Jansen, B. J., Booth, D. L., and Spink, A. (2008). Determining the informational, navigational, and transactional intent of web queries. *Information Processing and Management*, 44(3):1251–1266.
- Jansen, B. J. and Spink, A. (2005). Analysis of document viewing patterns of web search engine users. In *Web mining: Applications and techniques*, pages 339–354.
- Jansen, B. J. and Spink, A. (2006). How are we searching the world wide web? a comparison of nine search engine transaction logs. *Information Processing and Management*, 42(1):248–263.

- Järvelin, K. and Kekäläinen, J. (2000). Ir evaluation methods for retrieving highly relevant documents. In *Proceedings of the 23rd ACM SIGIR*, pages 41–48.
- Järvelin, K. and Kekäläinen, J. (2002). Cumulated gain-based evaluation of ir techniques. *ACM Transactions on Information Systems*, 20(4):422–446.
- Joachims, T. (2002). Optimizing search engines using clickthrough data. In *Proceedings of the 8th ACM KDD*, pages 133–142.
- Joachims, T., Granka, L., Pan, B., Hembrooke, H., and Gay, G. (2005). Accurately interpreting clickthrough data as implicit feedback. In *Proceedings of the 28th ACM SIGIR*, pages 154–161.
- Joho, H. and Jose, J. M. (2006). A comparative study of the effectiveness of search result presentation on the web. In *Proceedings of the 28th ECIR*, pages 302–313.
- Jones, K. S., Walker, S., and Robertson, S. E. (2000). A probabilistic model of information retrieval: Development and comparative experiments. *Information Processing and Management*, 36(6):779–808.
- Jordan, C., Watters, C., and Gao, Q. (2006). Using controlled query generation to evaluate blind relevance feedback algorithms. In *Proceedings of the 6th ACM/IEEE-CS JCDL*, pages 286–295.
- Kaisser, M., Hearst, M. A., and Lowe, J. B. (2008). Improving search results quality by customizing summary lengths. In *Proceedings of the 46th ACL*, pages 701–709.
- Kammerer, Y. and Gerjets, P. (2010). How the interface design influences users' spontaneous trustworthiness evaluations of web search results: comparing a list and a grid interface. In *Proceedings of the Symposium on Eye-Tracking Research & Applications*, pages 299–306.
- Kando, N., Eguchi, K., and Kuriyama, K. (1999). Construction of a large scale test collection: Analysis of the test topics of the NTCIR-1. In *Proceedings of IPSJ Annual Meeting*, pages 3–107.

- Kanungo, T. and Orr, D. (2009). Predicting the readability of short web summaries. In *Proceedings of the 2nd ACM WSDM*, pages 202–211.
- Kato, M. P., Sakai, T., and Tanaka, K. (2012). Structured query suggestion for specialization and parallel movement: Effect on search behaviors. In *Proceedings of the 21st WWW*, pages 389–398.
- Kazai, G., Kamps, J., Koolen, M., and Milic-Frayling, N. (2011). Crowdsourcing for book search evaluation: Impact of hit design on comparative system ranking. In *Proceedings of the 34th ACM SIGIR*, pages 205–214.
- Keenan, S., Smeaton, A. F., and Keogh, G. (2001). The effect of pool depth on system evaluation in trec. *Journal of the Association for Information Science and Technology*, 52(7):570–574.
- Kelly, D. (2009). Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval*, 3(1–2):1–224.
- Kelly, D., Arguello, J., Edwards, A., and Wu, W.-C. (2015). Development and evaluation of search tasks for iir experiments using a cognitive complexity framework. In *Proceedings of the 1st ACM ICTIR*, pages 101–110.
- Kelly, D. and Azzopardi, L. (2015). How many results per page?: A study of SERP size, search behavior and user experience. In *Proceedings of the 38th ACM SIGIR*, pages 183–192.
- Kelly, D. and Gyllstrom, K. (2011). An examination of two delivery modes for interactive search system experiments: Remote and laboratory. In *Proceedings of the 29th ACM CHI*, pages 1531–1540.
- Kelly, D., Gyllstrom, K., and Bailey, E. W. (2009). A comparison of query and term suggestion features for interactive searching. In *Proceedings of the 32nd ACM SIGIR*, pages 371–378.

- Keskustalo, H., Järvelin, K., Pirkola, A., Sharma, T., and Lykke, M. (2009). Test collection-based IR evaluation needs extension toward sessions — A case of extremely short queries. In *Proceedings of the 5th AIRS*, pages 63–74.
- Khabsa, M., Crook, A., Awadallah, A. H., Zitouni, I., Anastasakos, T., and Williams, K. (2016). Learning to account for good abandonment in search success metrics. In *Proceedings of the 25th ACM CIKM*, pages 1893–1896.
- Kim, J., Thomas, P., Sankaranarayana, R., and Gedeon, T. (2012). Comparing scanning behaviour in web search on small and large screens. In *Proceedings of the 17th ADCS*, pages 25–30.
- Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., and Yoon, H.-J. (2014). Eye-tracking analysis of user behavior and performance in web search on large and small screens. *Journal of the Association for Information Science and Technology*.
- Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., and Yoon, H.-J. (2016). Pagination versus scrolling in mobile web search. In *Proceedings of the 25th ACM CIKM*, pages 751–760.
- Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., and Yoon, H.-J. (2017). What snippet size is needed in mobile web search? In *Proceedings of the 2nd ACM CHIIR*, pages 97–106.
- Kiseleva, J., Kamps, J., Nikulin, V., and Makarov, N. (2015). Behavioral dynamics from the serp’s perspective: What are failed serps and how to fix them? In *Proceedings of the 24th ACM CIKM*, pages 1561–1570.
- Koch, S., Bosch, H., Giereth, M., and Ertl, T. (2009). Iterative integration of visual insights during patent search and analysis. In *Visual Analytics Science and Technology*, pages 203–210.
- Kogut, C. A. (1990). Consumer search behavior and sunk costs. *Journal of Economic Behavior and Organization*, 14(3):381–392.

- Kraft, D. and Lee, T. (1979). Stopping rules and their effect on expected search length. *IPM*, 15(1):47 – 58.
- Krebs, J. (1973). *Behavioral Aspects of Predation*, pages 73–111. Springer US, Boston, MA.
- Krebs, J., Ryan, J., and Charnov, E. (1974). Hunting by expectation or optimal foraging? a study of patch use by chickadees. *Animal Behaviour*, 22, Part 4:95–964.
- Krebs, J. R., Stephens, D. W., Sutherland, W. J., and Myers, J. P. (1983). Perspectives in optimal foraging. *Perspectives in Ornithology*, pages 165—222.
- Krovetz, R. (1993). Viewing morphology as an inference process. In *Proceedings of the 16th ACM SIGIR*, pages 191–202.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.
- Kupiec, J., Pedersen, J., and Chen, F. (1995). A trainable document summarizer. In *Proceedings of the 18th ACM SIGIR*, pages 68–73.
- Lancaster, F. (1968). *Information retrieval systems: characteristics, testing, and evaluation*. Information sciences series. Wiley.
- Landauer, T., Egan, D., Remde, J., Lesk, M., Lochbaum, C., and Ketchum, D. (1993). Enhancing the usability of text through computer delivery and formative evaluation: the superbook project. *Hypertext: A psychological perspective*, pages 71–136.
- Leal-Bando, L., Scholer, F., and Turpin, A. (2015). Query-biased summary generation assisted by query expansion. *JASIST*, 66(5):961–979.
- Li, Q. and Chen, Y. P. (2010). Personalized text snippet extraction using statistical language models. *Pattern Recognition*, 43(1):378–386.
- Li, Y. and Hu, D. (2013). Interactive retrieval using simulated versus real work task situations: Differences in sub-facets of tasks and interaction performance. In *Proceedings of the 76th ASIS&T*, pages 41:1–41:10.

- Linden, G. (2006). *Marissa mayer at web 2.0*. <http://glinden.blogspot.com/2006/11/marissa-mayer-at-web-20.html> (retrieved on August 10th, 2018).
- Lo, R. T.-W., He, B., and Ounis, I. (2005). Automatically building a stopword list for an information retrieval system. In *Proceedings of the 5th Dutch-Belgian IR Workshop*, pages 17–24.
- Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., and Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, 59(7):1041–1052.
- Loumakis, F., Stumpf, S., and Grayson, D. (2011). This image smells good: Effects of image info. scent in search engine results pages. In *Proceedings of the 20th ACM CIKM*, pages 475–484.
- Lovins, J. B. (1968). Development of a stemming algorithm. *Mechanical Translation and Computational Linguistics*, 11:22–31.
- Luhn, H. P. (1957). A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1(4):309–317.
- Luo, J., Zhang, S., Dong, X., and Yang, H. (2015). *Proceedings of the 37th ECIR*, chapter Designing States, Actions, and Rewards for Using POMDP in Session Search, pages 526–537.
- Luo, J., Zhang, S., and Yang, H. (2014). Win-win search: Dual-agent stochastic game in session search. In *Proceedings of the 37th ACM SIGIR*, pages 587–596.
- Manning, C. D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*.
- Mansourian, Y. and Ford, N. (2007). Search persistence and failure on the web: a “bounded rationality” and “satisficing” analysis. *Journal of Documentation*, 63(5):680–701.
- March, J. G. (1994). *Primer on decision making: How decisions happen*. Simon and Schuster.

- Marchionini, G. (1995). *Information Seeking in Electronic Environments*. Cambridge University Press.
- Marchionini, G., Dwiggins, S., Katz, A., and Lin, X. (1993). Information seeking in full-text end-user-oriented search systems: The roles of domain and search expertise. *Library and Information Science Research*, 15(1):35–69.
- Marcos, M.-C., Gavin, F., and Arapakis, I. (2015). Effect of snippets on user experience in web search. In *Proceedings of the 16th HCI*, pages 47:1–47:8.
- Marshall, C. C. and Shipman, F. M. (1997). Spatial hypertext and the practice of information triage. In *Proc. 8th ACM Hypertext*, pages 124–133.
- Maxwell, D. (2016). Building realistic simulations for interactive information retrieval. In *Proceedings of the 1st ACM CHIIR*, pages 357–359.
- Maxwell, D. and Azzopardi, L. (2014). Stuck in traffic: How temporal delays affect search behaviour. In *Proceedings of the 5th IliX*, pages 155–164.
- Maxwell, D. and Azzopardi, L. (2016a). Agents, simulated users and humans: An analysis of performance and behaviour. In *Proceedings of the 25th ACM CIKM*, pages 731–740.
- Maxwell, D. and Azzopardi, L. (2016b). Simulating interactive information retrieval: Simiir: A framework for the simulation of interaction. In *Proceedings of the 39th ACM SIGIR*, pages 1141–1144.
- Maxwell, D. and Azzopardi, L. (2018). Information scent, searching and stopping: Modelling SERP level stopping behaviour. In *Proceedings of the 40th ECIR*, pages 210–222.
- Maxwell, D., Azzopardi, L., Järvelin, K., and Keskustalo, H. (2015a). An initial investigation into fixed and adaptive stopping strategies. In *Proceedings of the 38th ACM SIGIR*, pages 903–906.

- Maxwell, D., Azzopardi, L., Järvelin, K., and Keskustalo, H. (2015b). Searching and stopping: An analysis of stopping rules and strategies. In *Proceedings of the 24th ACM CIKM*, pages 313–322.
- Maxwell, D., Azzopardi, L., and Moshfeghi, Y. (2017). A study of snippet length and informativeness: Behaviour, performance and user experience. In *Proceedings of the 40th ACM SIGIR*, pages 135–144.
- Maxwell, D., Azzopardi, L., and Moshfeghi, Y. (2019). The impact of result diversification on search behaviour and performance. *Information Retrieval Journal*. In press.
- McBryan, O. A. (1994). GENVL and WWW: Tools for taming the web. In *Proceedings of the 1st WWW*.
- McDonald, J., Ogden, W., and Foltz, P. (1998). Interactive information retrieval using term relationship networks. *NIST Special Publication*, pages 379–384.
- McMinn, A. J. (2018). *Real-Time Event Detection using Twitter*. PhD thesis, University of Glasgow.
- McNair, J. N. (1982). Optimal giving-up times and the marginal value theorem. *The American Naturalist*, 119(4):511–529.
- McNamee, P. (2006). Exploring new languages with haircut at clef 2005. In *Proceedings of the 6th CLEF*, pages 155–164.
- Mitra, B. and Craswell, N. (2017). Neural models for information retrieval. *arXiv preprint arXiv:1705.01509*.
- Moffat, A., Thomas, P., and Scholer, F. (2013). Users versus models: What observation tells us about effectiveness metrics. In *Proceedings of the 22nd ACM CIKM*, pages 659–668.
- Moffat, A. and Zobel, J. (2008). Rank-biased precision for measurement of retrieval effectiveness. *ACM Transactions on Information Systems*, 27(1):2:1–2:27.

- Mooers, C. (1950). *The theory of digital handling of non-numerical information and its implications to machine economics*. Zator technical bulletin.
- Moore, G. (1965). Cramming more components onto integrated circuits. *Electronics*, 38(8):114.
- Muralidharan, A., Gyongyi, Z., and Chi, E. (2012). Social annotations in web search. In *Proceedings of the 21st ACM CHI*, pages 1085–1094.
- Navalpakkam, V., Jentzsch, L., Sayres, R., Ravi, S., Ahmed, A., and Smola, A. (2013). Measurement and modeling of eye-mouse behavior in the presence of nonlinear page layouts. In *Proceedings of the 22nd WWW*, pages 953–964.
- Nickles, K. (1995). *Judgment-based and reasoning-based stopping rules in decision making under uncertainty*. PhD thesis, University of Minnesota.
- Ofoghi, B., Yearwood, J., and Ghosh, R. (2006). A semantic approach to boost passage retrieval effectiveness for question answering. In *Proceedings of the 29th ACSC*, pages 95–101.
- Olston, C. and Chi, E. (2003). Scenttrails: Integrating browsing and searching on the web. *ACM Transactions on Computer-Human Interactions*, 10(3).
- Ong, K., Järvelin, K., Sanderson, M., and Scholer, F. (2017). Using information scent to understand mobile and desktop web search behavior. In *Proceedings of the 40th ACM SIGIR*, pages 295–304.
- Oulasvirta, A., Hukkinen, J., and Schwartz, B. (2009). When more is less: The paradox of choice in search engine use. In *Proceedings of the 32nd ACM SIGIR*, pages 516–523.
- Over, P. (1998). Trec-6 interactive track report. pages 73–82.
- Over, P. (2001). The trec interactive track: an annotated bibliography. *Information Processing and Management*, 37(3):369–381.

- Pääkkönen, T., Järvelin, K., Kekäläinen, J., Keskustalo, H., Baskaya, F., Maxwell, D., and Azzopardi, L. (2015). Exploring behavioral dimensions in session effectiveness. In *Proceedings of the 6th CLEF*, pages 178–189.
- Paek, T., Dumais, S., and Logan, R. (2004). Wavelens: A new view onto internet search results. In *Proceedings of the 22nd ACM CHI*, pages 727–734.
- Pedersen, J., Cutting, D., Tukey, J., et al. (1991). Snippet search: A single phrase approach to text access. In *Proceedings of the 1991 Joint Statistical Meetings*.
- Perkins, D. N., Allen, R., and Hafner, J. (1983). Difficulties in everyday reasoning. *Thinking: The expanding frontier*, pages 177–189.
- Peters, C. and Braschler, M. (2001). European Research Letter: Cross-language System Evaluation: The CLEF Campaigns. *Journal of the Association for Information Science and Technology*, 52(12):1067–1072.
- Pirolli, P. (2007). *Information Foraging Theory: Adaptive interaction with information*. Human Technology Interaction Series. Oxford University Press, USA.
- Pirolli, P. and Card, S. (1995). Information foraging in information access environments. In *Proc. 13th ACM SIGCHI*, pages 51–58.
- Pirolli, P. and Card, S. K. (1999). Information foraging. *Psychological Review*, 106:643–675.
- Pirolli, P., Schank, P., Hearst, M., and Diehl, C. (1996). Scatter/gather browsing communicates the topic structure of a very large text collection. In *Proceedings of the 14th ACM CHI*, pages 213–220.
- Pitts, M. G. and Browne, G. J. (2004). Stopping behavior of systems analysts during information requirements elicitation. *Journal of Management Information Systems*, 21(1):203–226.
- Pitz, G., Reinhold, H., and Geller, E. S. (1969). Strategies of information seeking in deferred decision making. 4:1–19.

- Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3):130–137.
- Prabha, C., Connaway, L., Olszewski, L., and Jenkins, L. (2007). What is enough? Satisficing information needs. *Journal of Documentation*, 63(1):74–89.
- Reisberg, D. (1997). *Cognition: Exploring the science of the mind*. WW Norton & Co.
- Resnick, M. L., Maldonado, C., Santos, J. M., and Lergier, R. (2001). Modeling on-line search behavior using alternative output structures. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 45, pages 1166–1170.
- Robertson, S. (2008). On the history of evaluation in ir. *Journal of Information Science*, 34(4):439–456.
- Robertson, S., Walker, S., Jones, S., Hancock-Beaulieu, M. M., and Gatford, M. (1995). Okapi at trec-3. In *Overview of the Third Text REtrieval Conference (TREC-3)*, page 109–126.
- Robertson, S. and Zaragoza, H. (2009). The probabilistic relevance framework: Bm25 and beyond. *Found. Trends Inf. Retr.*, 3(4):333–389.
- Robertson, S. E. (1977). The probability ranking principle in ir. *Journal of Documentation*, 33(4):294–304.
- Rose, D. E. and Levinson, D. (2004). Understanding user goals in web search. In *Proceedings of the 13th WWW*, pages 13–19.
- Rose, D. E., Orr, D., and Kantamneni, R. G. P. (2007). Summary attributes and perceived search quality. In *Proceedings of the 16th WWW*, pages 1201–1202.
- Russell, D. M., Stefik, M. J., Pirolli, P., and Card, S. K. (1993). The cost structure of sense-making. In *Proceedings of the 11th ACM CHI*, pages 269–276.
- Ruthven, I. (2001). *Abduction, Explanation and Relevance Feedback*. PhD thesis, University of Glasgow.

- Ruthven, I. (2008). Interactive information retrieval. *Annual Review of Information Science and Technology*, 42(1):43–91.
- Saad, G. and Russo, J. (1996). Stopping criteria in sequential choice. *Organizational Behavior and Human Decision Processes*, 67(3):258 – 270.
- Salton, G., Wong, A., and Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620.
- Sanderson, M. (2010). Test collection based evaluation of information retrieval systems. *Foundations and Trends in Information Retrieval*, 4(4):247–375.
- Sanderson, M. and Croft, W. B. (2012). The history of information retrieval research. *Proceedings of the IEEE*, 100:1444–1451.
- Sandstrom, P. E. (1994). An optimal foraging approach to information seeking and use. *The Library Quarterly: Information, Community, Policy*, 64(4):414–449.
- Santos, R. L., Macdonald, C., and Ounis, I. (2010). Exploiting query reformulations for web search result diversification. In *Proceedings of the 19th WWW*, pages 881–890.
- Savenkov, D., Braslavski, P., and Lebedev, M. (2011). Search snippet evaluation at yandex: lessons learned and future directions. *Multilingual and Multimodal Information Access Evaluation*, pages 14–25.
- Schank, R. C. and Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Psychology Press.
- Schwartz, B. (2005). *The Paradox of Choice: Why More Is Less*. Harper Perennial.
- Shafir, E. and Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive psychology*, 24(4):449–474.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

- Simon, H. A. (1971). Decision making and organizational design. *Organizational Theory*, pages 189–212.
- Simon, H. A. (1996). *The sciences of the artificial*. MIT press.
- Smith, G. F., Benson, P. G., and Curley, S. P. (1991). Belief, knowledge, and uncertainty: A cognitive perspective on subjective probability. *Organizational Behavior and Human Decision Processes*, 48(2):291–321.
- Smucker, M. (2011). An analysis of user strategies for examining and processing ranked lists of documents. In *Proceedings of the 5th HCIR*.
- Smucker, M., Guo, X., and Toulis, A. (2014). Mouse movement during relevance judging: Implications for determining user attention. In *Proceedings of the 37th ACM SIGIR*, pages 979–982.
- Smucker, M. D. and Clarke, C. L. (2012). Time-based calibration of effectiveness measures. In *Proceedings of the 35th ACM SIGIR*, pages 95–104.
- Soper, H. (1918). Means for compiling tabular and statistical data. U.S. Patent US00135169231–1920.
- Spärck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1):11–21.
- Spirin, N. V., Kotov, A. S., Karahalios, K. G., Mladenov, V., and Izhutov, P. A. (2016). A comparative study of query-biased and non-redundant snippets for structured search on mobile devices. In *Proceedings of the 25th ACM CIKM*, pages 2389–2394.
- Spool, J. and Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. In *Proc. 19th ACM CHI Extended Abstracts*, pages 285–286.
- Stephens, D. and Krebs, J. (1986). *Foraging Theory*. Monographs in Behavior and Ecology. Princeton University Press.

- Su, L. T. (1992). Evaluation measures for interactive information retrieval. *Information Processing and Management*, 28(4):503–516. Special Issue: Evaluation Issues in Information Retrieval.
- Sundar, S., Knobloch-Westerwick, S., and Hastall, M. (2007). News cues: Info. scent and cognitive heuristics. *Journal of the Association for Information Science Technology*, 58(3):366–378.
- Svore, K. M., Teevan, J., Dumais, S. T., and Kulkarni, A. (2012). Creating temporally dynamic web search snippets. In *Proceedings of the 35th ACM SIGIR*, pages 1045–1046.
- Syed, R. and Collins-Thompson, K. (2017). Retrieval algorithms optimized for human learning. In *Proceedings of the 40th ACM SIGIR*, pages 555–564.
- Teevan, J., Cutrell, E., Fisher, D., Drucker, S. M., Ramos, G., André, P., and Hu, C. (2009). Visual snippets: Summarizing web pages for search and revisitation. In *Proceedings of the 27th ACM CHI*, pages 2023–2032.
- Thomas, P., Moffat, A., Bailey, P., and Scholer, F. (2014). Modeling decision points in user search behavior. In *Proceedings of the 5th IIX*, pages 239–242.
- Tocher, K. (1963). *The art of simulation*. Electrical engineering series. English Universities Press.
- Tombros, A. and Sanderson, M. (1998). Advantages of query biased summaries in information retrieval. In *Proceedings of the 21st ACM SIGIR*, pages 2–10.
- Toms, E. G. and Freund, L. (2009). Predicting stopping behaviour: A preliminary analysis. In *Proceedings of the 32nd ACM SIGIR*, pages 750–751.
- Tran, V., Maxwell, D., Fuhr, N., and Azzopardi, L. (2017). Personalised search time prediction using markov chains. In *Proceedings of the 3rd ACM ICTIR*, pages 237–240.
- Turpin, A., Scholer, F., Jarvelin, K., Wu, M., and Culpepper, J. S. (2009). Including summaries in system evaluation. In *Proceedings of the 32nd ACM SIGIR*, pages 508–515.

- Turpin, A., Tsegay, Y., Hawking, D., and Williams, H. E. (2007). Fast generation of result snippets in web search. In *Proceedings of the 30th ACM SIGIR*, pages 127–134.
- Umemoto, K., Yamamoto, T., and Tanaka, K. (2016). Scentbar: A query suggestion interface visualizing the amount of missed relevant information for intrinsically diverse search. In *Proceedings of the 39th ACM SIGIR*, pages 405–414.
- van Rijsbergen, C. (1979). *Information Retrieval*. Butterworth-Heinemann.
- Vaughan, L. (2004). New measurements for search engine evaluation proposed and tested. *Information Processing and Management*, 40(4):677–691.
- Veerasamy, A. and Belkin, N. J. (1996). Evaluation of a tool for visualization of information retrieval results. In *Proceedings of the 19th ACM SIGIR*, pages 85–92.
- Veerasamy, A. and Heikes, R. (1997). Effectiveness of a graphical display of retrieval results. In *Proceedings of the 20th ACM SIGIR*, pages 236–245.
- Verberne, S., Sappelli, M., Järvelin, K., and Kraaij, W. (2015). User simulations for interactive search: Evaluating personalized query suggestion. In *Proceedings of the 37th ECIR*, volume 9022, pages 678–690.
- Villa, R., Cantador, I., Joho, H., and Jose, J. M. (2009). An aspectual interface for supporting complex search tasks. In *Proceedings of the 32nd ACM SIGIR*, pages 379–386.
- Voorhees, E. (2006). Overview of the trec 2005 robust retrieval track. In *Proceedings of TREC-14*.
- Voorhees, E. M. (2002). The philosophy of information retrieval evaluation. In *Proceedings of the 7th CLEF Initiative*, pages 355–370.
- Voorhees, E. M. and Harman, D. K. (2005). *TREC: Experiment and Evaluation in Information Retrieval*.
- Wald, A. (1948). Sequential analysis. *Social Forces*, 27(2):170–171.

- White, R. W. and Dumais, S. T. (2009). Characterizing and predicting search engine switching behavior. In *Proceedings of the 18th ACM CIKM*, pages 87–96.
- White, R. W., Jose, J. M., and Ruthven, I. (2003). A task-oriented study on the influencing effects of query-biased summarisation in web searching. *Information Processing and Management*, 39(5):707–733.
- Wilkie, C. and Azzopardi, L. (2017). Algorithmic bias: Do good systems make relevant documents more retrievable? In *Proceedings of the 26th ACM CIKM*, pages 2375–2378.
- Wilson, M. L., Kules, B., Schraefel, M., and Shneiderman, B. (2010). From keyword search to exploration: Designing future search interfaces for the web. *Foundations and Trends in Web Science*, 2(1):1–97.
- Woodruff, A., Rosenholtz, R., Morrison, J. B., Faulring, A., and Pirolli, P. (2002). A comparison of the use of text summaries, plain thumbnails, and enhanced thumbnails for web search tasks. *Journal of the American Society for Information Science and Technology*, 53(2):172–185.
- Wu, W.-C. (2012). How far will you go?: Using need for closure and information scent to model search stopping behavior. In *Proceedings of the 4th IiX*, pages 328–328.
- Wu, W.-C. and Kelly, D. (2014). Online search stopping behaviors: An investigation of query abandonment and task stopping. *Proceedings of the 77^{ASIST}*, 51(1):1–10.
- Wu, W.-C., Kelly, D., and Sud, A. (2014). Using information scent and need for cognition to understand online search behavior. In *Proceedings of the 37th ACM SIGIR*, pages 557–566.
- Yang, P. and Fang, H. (2017). Can short queries be even shorter? In *Proceedings of the 40th ACM SIGIR*, pages 43–50.
- Yates, J. (1990). Judgment and decision making. *Journal of Behavioral Decision Making*, 4(1):76–78.

- Yates, J. and Carlson, B. (1982). Toward a representational theory of decision making. *Ann Arbor, MI: University of Michigan, Working Paper.*
- Yilmaz, E., Shokouhi, M., Craswell, N., and Robertson, S. (2010). Expected browsing utility for web search evaluation. In *Proceedings of the 19th ACM CIKM*, pages 1561–1564.
- Zach, L. (2005). When is “enough” enough? modeling the information-seeking and stopping behavior of senior arts administrators: Research articles. *Journal of the Association for Information Science and Technology*, 56(1):23–35.
- Zhang, F., Liu, Y., Li, X., Zhang, M., Xu, Y., and Ma, S. (2017a). Evaluating web search with a bejeweled player model. In *Proceedings of the 40th ACM SIGIR*, pages 425–434.
- Zhang, Y., Liu, X., and Zhai, C. X. (2017b). Information retrieval evaluation as search simulation: A general formal framework for ir evaluation. In *Proceedings of the 3rd ACM ICTIR*, pages 193–200.
- Zhang, Y., Park, L. A. F., and Moffat, A. (2010). Click-based evidence for decaying weight distributions in search effectiveness metrics. *Information Retrieval*, 13(1):46–69.
- Zipf, G. (1949). *Human behavior and the principle of least effort: an introduction to human ecology.*
- Zuccon, G., Leelanupab, T., Whiting, S., Yilmaz, E., Jose, J. M., and Azzopardi, L. (2013). Crowdsourcing interactions: Using crowdsourcing for evaluating interactive information retrieval systems. *Information Retrieval*, 16(2):267–305.