

3 Bibliography

- [1] Adomavicius, G. and Tuzhilin, A. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*. 17, 6 (2005), 734–749.
- [2] Agrawal, R., Gollapudi, S., Kannan, A. and Kenthapadi, K. 2012. Data mining for improving textbooks. *SIGKDD Explorations Newsletter*. 13, 2 (2012), 7–19.
- [3] Agrawal, R., Gollapudi, S., Kenthapadi, K., Srivastava, N. and Velu, R. 2010. Enriching textbooks through data mining. *Proc. ACM Symposium on Computing for Development* (2010).
- [4] Agrawal, R., Imieliński, T. and Swami, A. 1993. Mining association rules between sets of items in large databases. *SIGMOD Record*. 22, 2 (1993), 207–216.
- [5] Ammar, W., Mulcaire, G., Tsvetkov, Y., Lample, G., Dyer, C. and Smith, N.A. 2016. Massively multilingual word embeddings. *arXiv:1602.01925*.
- [6] André, P., Schraefel, M. c., Teevan, J. and Dumais, S.T. 2009. Discovery is never by chance: Designing for (un)serendipity. *Proc. ACM Conference on Creativity and Cognition* (2009).
- [7] Bao, P., Hecht, B., Carton, S., Quaderi, M., Horn, M. and Gergle, D. 2012. Omnipedia: Bridging the Wikipedia language gap. *Proc. SIGCHI Conference on Human Factors in Computing Systems* (2012).
- [8] Bayliss, G. 2013. Exploring the cautionary attitude toward Wikipedia in higher education: Implications for higher education institutions. *New Review of Academic Librarianship*. 19, 1 (2013), 36–57.
- [9] Bilenko, M. and White, R.W. 2008. Mining the search trails of surfing crowds: Identifying relevant websites from user activity. *Proc. International Conference on the World Wide Web* (2008).
- [10] Boldi, P. and Monti, C. 2016. Cleansing Wikipedia categories using centrality. *Companion of the International Conference on the World Wide Web* (2016).
- [11] Brin, S. and Page, L. 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*. 30, 1 (1998), 107–117.
- [12] Chen, H. 2010. The perspectives of higher education faculty on Wikipedia. *The Electronic Library*. 28, 3 (2010), 361–373.
- [13] Chi, E.H., Pirolli, P., Chen, K. and Pitkow, J. 2001. Using information scent to model user information needs and actions and the Web. *Proc. SIGCHI Conference on Human Factors in Computing Systems* (2001).
- [14] Chierichetti, F., Kumar, R., Raghavan, P. and Sarlos, T. 2012. Are Web users really Markovian? *Proc. International Conference on the World Wide Web* (2012).
- [15] Clauset, A. and Moore, C. 2003. How do networks become navigable? *arXiv:cond-mat/0309415*.
- [16] Cosley, D., Frankowski, D., Terveen, L. and Riedl, J. 2007. SuggestBot: Using intelligent task routing to help people find work in Wikipedia. *Proc. International Conference on Intelligent User Interfaces* (2007).
- [17] Dimitrov, D., Singer, P., Lemmerich, F. and Strohmaier, M. 2016. Visual positions of links and clicks on Wikipedia. *Companion of the International Conference on the World Wide Web* (2016).
- [18] Dodds, P.S., Muhamad, R. and Watts, D.J. 2003. An experimental study of search in global social networks. *Science*. 301, 5634 (2003), 827–829.

- [19] Downey, D., Dumais, S., Liebling, D. and Horvitz, E. 2008. Understanding the relationship between searchers' queries and information goals. *Proc. ACM Conference on Information and Knowledge Management* (2008).
- [20] Eom, Y.-H. and Shepelyansky, D.L. 2013. Highlighting entanglement of cultures via ranking of multilingual Wikipedia articles. *PloS one*. 8, 10 (2013), e74554.
- [21] Flati, T., Vannella, D., Pasini, T. and Navigli, R. 2016. MultiWiBi: The multilingual Wikipedia bitaxonomy project. *Artificial Intelligence*. 241, (2016), 66–102.
- [22] Gibson, D., Kumar, R. and Tomkins, A. 2005. Discovering large dense subgraphs in massive graphs. *Proc. International Conference on Very Large Data Bases* (2005).
- [23] Giles, J. 2005. Internet encyclopaedias go head to head. *Nature*. 438, 7070 (2005), 900–901.
- [24] Guo, Q., Kulkarni, C., Kittur, A., Bigham, J.P. and Brunskill, E. 2016. Questimator: Generating knowledge assessments for arbitrary topics. *Proc. International Joint Conference on Artificial Intelligence* (2016).
- [25] Gupta, A., Piccinno, F., Kozhevnikov, M., Pasca, M. and Pighin, D. 2016. Revisiting taxonomy induction over Wikipedia. *Proc. International Conference on Computational Linguistics* (2016).
- [26] Halfaker, A., Geiger, R.S. and Morgan, J.T. 2013. The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline. *American Behavioral Scientist*. 57, 5 (2013), 664–688.
- [27] Hazari, S., North, A. and Moreland, D. 2009. Investigating pedagogical value of wiki technology. *Journal of Information Systems Education*. 20, 2 (2009), 187–198.
- [28] Helic, D., Strohmaier, M., Granitzer, M. and Scherer, R. 2013. Models of human navigation in information networks based on decentralized search. *Proc. ACM Conference on Hypertext and Social Media* (2013).
- [29] Hochreiter, S. and Schmidhuber, J. 1997. Long short-term memory. *Neural computation*. 9, 8 (1997), 1735–1780.
- [30] Kleinberg, J. 2000. The small-world phenomenon: An algorithmic perspective. *Proc. Annual ACM Symposium on Theory of Computing* (2000).
- [31] Kumar, S., West, R. and Leskovec, J. 2016. Disinformation on the Web: Impact, characteristics, and detection of Wikipedia hoaxes. *Proc. International Conference on the World Wide Web* (2016).
- [32] Laufer, P., Wagner, C., Flöck, F. and Strohmaier, M. 2015. Mining cross-cultural relations from Wikipedia: A study of 31 European food cultures. *Proc. ACM Web Science Conference* (2015).
- [33] Lehmann, J., Müller-Birn, C., Laniado, D., Lalmas, M. and Kaltenbrunner, A. 2014. Reader preferences and behavior on Wikipedia. *Proc. ACM Conference on Hypertext and Social Media* (2014).
- [34] Leskovec, J., Kleinberg, J. and Faloutsos, C. 2005. Graphs over time: Densification laws, shrinking diameters and possible explanations. *Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2005).
- [35] Liben-Nowell, D., Novak, J., Kumar, R., Raghavan, P. and Tomkins, A. 2005. Geographic routing in social networks. *Proceedings of the National Academy of Sciences*. 102, 33 (2005), 11623–11628.
- [36] Li, Z. and Tian, J. 2003. Testing the suitability of Markov chains as Web usage models. *Proc. Annual International Computer Software and Applications Conference* (2003).
- [37] Mahdian, M., West, R. and Zia, L. 2015. Automatic bilingual dictionary generation using Wikipedia. *ICWSM Workshop on Wikipedia, a Social Media* (2015).
- [38] McDaniel, M.A., Agarwal, P.K. and Huelser, B.J. 2011. Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*. 103, 2 (2011), 399–414.
- [39] McDowell, Z. 2017. Student learning outcomes using Wikipedia-based assignments: Fall 2016 research report.

Wikimedia Commons. (2017).

- [40] de Melo, G. and Weikum, G. 2010. MENTA: Inducing multilingual taxonomies from Wikipedia. *Proc. ACM International Conference on Information and Knowledge Management* (2010).
- [41] Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J. 2013. Distributed representations of words and phrases and their compositionality. *Proc. Advances in Neural Information Processing Systems* (2013).
- [42] Milgram, S. 1967. The small world problem. *Psychology Today*. 2, 1 (1967), 60–67.
- [43] Moy, C.L., Locke, J.R., Coppola, B.P. and McNeil, A.J. 2010. Improving science education and understanding through editing Wikipedia. *Journal of Chemical Education*. 87, 11 (2010), 1159–1162.
- [44] Navigli, R. and Ponzetto, S.P. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*. 193, (2012), 217–250.
- [45] Nemhauser, G.L., Wolsey, L.A. and Fisher, M.L. 1978. An analysis of approximations for maximizing submodular set functions -- I. *Mathematical Programming*. 14, 1 (1978), 265–294.
- [46] Nielsen, F.Å. 2007. Scientific citations in Wikipedia. *arXiv:0705.2106*.
- [47] Ni, X., Sun, J.-T., Hu, J. and Chen, Z. 2009. Mining multilingual topics from Wikipedia. *Proc. International Conference on the World wide web* (2009).
- [48] Nothman, J., Ringland, N., Radford, W., Murphy, T. and Curran, J.R. 2013. Learning multilingual named entity recognition from Wikipedia. *Artificial Intelligence*. 194, (2013), 151–175.
- [49] Okoli, C., Mehdi, M., Mesgari, M., Nielsen, F.Å. and Lanamäki, A. 2014. Wikipedia in the eyes of its beholders: A systematic review of scholarly research on Wikipedia readers and readership. *Journal of the Association for Information Science and Technology*. 65, 12 (2014), 2381–2403.
- [50] Paranjape, A., West, R., Zia, L. and Leskovec, J. 2016. Improving website hyperlink structure using server logs. *Proc. International ACM Conference on Web Search and Data Mining*. (2016).
- [51] Pirolli, P. 2007. *Information Foraging Theory: Adaptive Interaction with Information*. Oxford University Press.
- [52] Pirolli, P. 2005. Rational analyses of information foraging on the Web. *Cognitive Science*. 29, 3 (2005), 343–373.
- [53] Pirolli, P. and Card, S.K. 1999. Information foraging. *Psychological Review*. 106, 4 (1999), 643–675.
- [54] Platt, J.C., Toutanova, K. and Yih, W.-T. 2010. Translingual document representations from discriminative projections. *Proc. Conference on Empirical Methods in Natural Language Processing* (2010).
- [55] Potthast, M., Stein, B. and Anderka, M. 2008. A Wikipedia-based multilingual retrieval model. *Advances in Information Retrieval*. C. Macdonald, I. Ounis, V. Plachouras, I. Ruthven, and R.W. White, eds. Springer Berlin Heidelberg. 522–530.
- [56] Rajpurkar, P., Zhang, J., Lopyrev, K. and Liang, P. 2016. SQuAD: 100,000+ questions for machine comprehension of text. *arXiv:1606.05250*.
- [57] Sahin, C.S., Caceres, R.S., Oselio, B. and Campbell, W.M. 2017. Consistent alignment of word embedding models. *arXiv:1702.07680*.
- [58] Samoilenko, A., Karimi, F., Edler, D., Kunegis, J. and Strohmaier, M. 2016. Linguistic neighbourhoods: Explaining cultural borders on Wikipedia through multilingual co-editing activity. *EPJ Data Science*. 5, 1 (2016), 9.
- [59] Sarukkai, R.R. 2000. Link prediction and path analysis using Markov chains. *Computer Networks*. 33, 1 (2000), 377–386.
- [60] Schwarzer, M., Schubotz, M., Meuschke, N., Breiting, C., Markl, V. and Gipp, B. 2016. Evaluating link-based recommendations for Wikipedia. *Proc. ACM/IEEE Joint Conference on Digital Libraries* (2016).

- [61] Sen, R. and Hansen, M.H. 2003. Predicting Web users' next access based on log data. *Journal of Computational and Graphical Statistics*. 12, 1 (2003), 143–155.
- [62] Şimşek, Ö. and Jensen, D. 2008. Navigating networks by using homophily and degree. *Proceedings of the National Academy of Sciences*. 105, 35 (2008), 12758–12762.
- [63] Singer, P., Helic, D., Hotho, A. and Strohmaier, M. 2015. HypTrails: A Bayesian approach for comparing hypotheses about human trails on the Web. *Proc. International Conference on the World Wide Web* (2015).
- [64] Singer, P., Lemmerich, F., West, R., Zia, L., Wulczyn, E., Strohmaier, M. and Leskovec, J. 2017. Why we read Wikipedia. *Proc. International Conference on the World Wide Web* (2017).
- [65] Singla, A., White, R. and Huang, J. 2010. Studying trailfinding algorithms for enhanced Web search. *Proc. International ACM SIGIR Conference on Research and Development in Information Retrieval* (2010).
- [66] Talukdar, P.P. and Cohen, W.W. 2012. Crowdsourced comprehension: Predicting prerequisite structure in Wikipedia. *Proc. Workshop on Building Educational Applications Using NLP* (2012).
- [67] Teevan, J., Alvarado, C., Ackerman, M.S. and Karger, D.R. 2004. The perfect search engine is not enough: A study of orienteering behavior in directed search. *Proc. SIGCHI Conference on Human Factors in Computing Systems* (2004).
- [68] West, R. 2016. *Human Navigation of Information Networks*. Stanford University.
- [69] West, R., Gabrilovich, E., Murphy, K., Sun, S., Gupta, R. and Lin, D. 2014. Knowledge base completion via search-based question answering. *Proc. International Conference on the World Wide Web* (2014).
- [70] West, R. and Leskovec, J. 2012. Automatic versus human navigation in information networks. *Proc. International Conference on Weblogs and Social Media* (2012).
- [71] West, R. and Leskovec, J. 2012. Human wayfinding in information networks. *Proc. International Conference on the World Wide Web* (2012).
- [72] West, R., Paranjape, A. and Leskovec, J. 2015. Mining missing hyperlinks from human navigation traces: A case study of Wikipedia. *Proc. International Conference on the World Wide Web* (2015).
- [73] West, R. and Pfeffer, J. 2017. Armed conflicts in online news: A multilingual study. *Proc. International Conference on Web and Social Media* (2017).
- [74] West, R., Pineau, J. and Precup, D. 2009. Wikispeedia: An online game for inferring semantic distances between concepts. *Proc. International Joint Conference on Artificial Intelligence* (2009).
- [75] West, R., Precup, D. and Pineau, J. 2010. Automatically suggesting topics for augmenting text documents. *Proc. ACM Conference on Information and Knowledge Management* (2010).
- [76] West, R., Precup, D. and Pineau, J. 2009. Completing Wikipedia's hyperlink structure through dimensionality reduction. *Proc. ACM Conference on Information and Knowledge Management* (2009).
- [77] West, R., Weber, I. and Castillo, C. 2012. Drawing a data-driven portrait of Wikipedia editors. *Proc. Annual International Symposium on Wikis and Open Collaboration* (2012).
- [78] West, R., White, R.W. and Horvitz, E. 2013. From cookies to cooks: Insights on dietary patterns via analysis of Web usage logs. *Proc. International Conference on the World Wide Web* (2013).
- [79] West, R., White, R.W. and Horvitz, E. 2013. Here and there: Goals, activities, and predictions about location from geotagged queries. *Proc. International ACM SIGIR Conference on Research and Development in Information Retrieval* (2013).
- [80] White, R.W. and Huang, J. 2010. Assessing the scenic route: Measuring the value of search trails in Web logs. *Proc.*

International ACM SIGIR Conference on Research and Development in Information Retrieval (2010).

- [81] White, R.W. and Singla, A. 2011. Finding our way on the Web: Exploring the role of waypoints in search interaction. *Proc. International Conference on the World Wide Web* (2011).
- [82] Wikipedia clickstream: 2015. <https://doi.org/10.6084/m9.figshare.1305770.v22>.
- [83] Wikipedia navigation vectors: 2017. <http://dx.doi.org/10.6084/m9.figshare.3146878.v6>. Accessed: 2017-10-01.
- [84] Wulczyn, E., West, R., Zia, L. and Leskovec, J. 2016. Growing Wikipedia across languages via recommendation. *Proc. International Conference on the World Wide Web* (2016).
- [85] Yasuda, K. and Sumita, E. 2008. Method for building sentence-aligned corpus from Wikipedia. *Proc. AAAI Workshop on Wikipedia and Artificial Intelligence* (2008).
- [86] Yih, W.-T., Toutanova, K., Platt, J.C. and Meek, C. 2011. Learning discriminative projections for text similarity measures. *Proc. Conference on Computational Natural Language Learning* (2011).
- [87] Yuan, X., Wang, T., Gulcehre, C., Sordoni, A., Bachman, P., Subramanian, S., Zhang, S. and Trischler, A. 2017. Machine comprehension by text-to-text neural question generation. *arXiv:1705.02012*.
- [88] Zukerman, I., Albrecht, D.W. and Nicholson, A.E. 1999. Predicting users' requests on the WWW. *Proc. International Conference on User Modeling* (1999).