

Initial Article Set resulting from unstructured Search

- [1] K. Boland, D. Ritze, K. Eckert, and B. Mathiak, “Identifying references to datasets in publications,” in *International Conference on Theory and Practice of Digital Libraries*. Springer, 2012, pp. 150–161.
- [2] B. Ghavimi, P. Mayr, C. Lange, S. Vahdati, and S. Auer, “A semi-automatic approach for detecting dataset references in social science texts,” *Information Services & Use*, vol. 36, no. 3-4, pp. 171–187, 2016.
- [3] B. Ghavimi, P. Mayr, S. Vahdati, and C. Lange, “Identifying and improving dataset references in social sciences full texts,” *arXiv preprint arXiv:1603.01774*, 2016.
- [4] M. Lu, S. Bangalore, G. Cormode, M. Hadjieleftheriou, and D. Srivastava, “A dataset search engine for the research document corpus,” in *Data Engineering (ICDE), 2012 IEEE 28th International Conference on*. IEEE, 2012, pp. 1237–1240.
- [5] H. Mooney and M. P. Newton, “The anatomy of a data citation: Discovery, reuse, and credit.” *Journal of Librarianship & Scholarly Communication*, vol. 1, no. 1, 2012.
- [6] H. A. Piwowar, J. D. Carlson, and T. J. Vision, “Beginning to track 1000 datasets from public repositories into the published literature,” *Proceedings of the Association for Information Science and Technology*, vol. 48, no. 1, pp. 1–4, 2011.
- [7] A. Singhal and J. Srivastava, “Data extract: Mining context from the web for dataset extraction,” *International Journal of Machine Learning and Computing*, vol. 3, no. 2, p. 219, 2013.
- [8] Q. Zhang, Q. Cheng, Y. Huang, and W. Lu, “A bootstrapping-based method to automatically identify data-usage statements in publications,” *Journal of Data and Information Science*, vol. 1, no. 1, pp. 69–85, 2016.
- [9] M. Zhao, E. Yan, and K. Li, “Data set mentions and citations: A content analysis of full-text publications,” *Journal of the Association for Information Science and Technology*, 2018.
- [10] J. Howison and J. Bullard, “Software in the scientific literature: Problems with seeing, finding, and using software mentioned in the biology literature,” *Journal of the Association for Information Science and Technology*, vol. 67, no. 9, pp. 2137–2155, may 2015.
- [11] K. Li and E. Yan, “Co-mention network of r packages: Scientific impact and clustering structure,” *Journal of Informetrics*, vol. 12, no. 1, pp. 87–100, feb 2018.

- [12] K. Li, E. Yan, and Y. Feng, “How is r cited in research outputs? structure, impacts, and citation standard,” *Journal of Informetrics*, vol. 11, no. 4, pp. 989–1002, nov 2017.
- [13] X. Pan, E. Yan, and W. Hua, “Disciplinary differences of software use and impact in scientific literature,” *Scientometrics*, vol. 109, no. 3, pp. 1593–1610, sep 2016.