Bibliometric-enhanced information retrieval: preface

Guillaume Cabanac¹ · Ingo Frommholz² · Philipp Mayr³

Received: 16 July 2018 / Published online: 21 July 2018
© Akadémiai Kiadó, Budapest, Hungary 2018

The special issue on Bibliometric-enhanced Information Retrieval presents works at the crossroad between Bibliometrics and Information Retrieval, two domains closely related on the map of Information Science (Fig. 1). With the surge of “scholarly big data” (Giles 2013) Bibliometrics and Information Retrieval have seen a recent renaissance that resulted in a recent special issue on “Combining bibliometrics and information retrieval” in *Scientometrics* in 2015 (Mayr and Scharnhorst 2015) and a successful workshop series held at leading conferences such as SIGIR, JCDL, and ECIR and covering topics relevant not just to *Scientometrics*/Bibliometrics and Information Retrieval, but also to Natural Language Processing, Machine Learning, and Digital Libraries.

This special issue was announced *via* an open call for papers. Among the eighteen submissions we received, most were extended workshop papers from the BIR workshop at ECIR 2017 held in Aberdeen, Scotland (Mayr et al. 2017b) and the BIRNDL workshop at SIGIR 2017 held in Tokyo, Japan (Mayr et al. 2017a). All submissions were evaluated by two to three referees with an overall balanced expertise in Bibliometrics and Information Retrieval.

This special issue eventually includes eight papers: three are extended papers presented at BIR and five are extended papers presented at BIRNDL. It covers a variety of topics including: Authorship identification and writing style analysis (Rexha et al. 2018); Locality sensitive hashing to measure the overlap bibliographic databases for paper indexing (Abdulhayoglu and Thijs 2018); Automatic discovery of cross-topic collaborations (Cagliero et al. 2018); Identification of problems and their solutions in scientific articles (Heffernan and Teufel 2018); Accurate identification of cited text spans (e.g., to

✉ Philipp Mayr
philipp.mayr@gesis.org

Guillaume Cabanac
guillaume.cabanac@univ-tlse3.fr

Ingo Frommholz
ifrommholz@acm.org

¹ Computer Science Department, IRIT UMR 5505 CNRS, University of Toulouse, 118 route de Narbonne, 31062 Toulouse Cedex 9, France

² University of Bedfordshire, Luton, LU1 3JU, UK

³ GESIS – Leibniz Institute for the Social Sciences, Cologne, Germany

Fig. 1 Structure of the Information Science field in 2006–2015 as seen from the first-author AKCA (author keyword coupling analysis). Reprinted from Yang et al. (2016, Fig. 2) with permission from Elsevier

generate summaries) (Ma et al. 2018); Detection of automatically generated sentences and texts within genuinely written papers (Tien and Labbé 2018); Citance-based retrieval and summarization using Information Retrieval and Machine Learning (Karimi et al. 2018); Analysis of search stratagems for interactive information retrieval in bibliographic databases (Kacem and Mayr 2018).

We hope the selection of papers in this special issue will be interesting and enjoyable for researchers coming from all relevant fields and provide a starting point for future explorations in the field.

There is an other special issue on “Bibliometric-enhanced Information Retrieval and Natural Language Processing for Digital Libraries” that was compiled after the first joint BIRNDL workshop held at JCDL 2016 in Newark, New Jersey, USA and is coming out in the “International Journal on Digital Libraries”. This special issue contains fourteen papers (Mayr et al. forthcoming).

Acknowledgements We wish to thank all those who contributed to this special issue: The researchers who contributed papers, the many reviewers who generously offered their time and expertise, the various people involved in publishing the issue, and the participants of the BIR and BIRNDL workshops. (Since 2016 we maintain the “Bibliometric-enhanced-IR Bibliography” https://github.com/PhilippMayr/Bibliometric-enhanced-IR_Bibliography/ that collects scientific papers which appear in collaboration with the BIR/ BIRNDL organizers.)

References

- Abdulhayoglu, M. A., & Thijs, B. (2018). Use of locality sensitive hashing (LSH) algorithm to match Web of Science and Scopus. *Scientometrics*. <https://doi.org/10.1007/s11192-017-2569-6>.
- Cagliero, L., Garza, P., Kavosifar, M. R., & Baralis, E. (2018). Discovering cross-topic collaborations among researchers by exploiting weighted association rules. *Scientometrics*. <https://doi.org/10.1007/s11192-018-2737-3>.

- Giles, C. L. (2013). Scholarly big data. In: *CIKM'13: Proceedings of the 22nd ACM international conference on conference on information and knowledge management* (p. 1). ACM, New York, NY. <https://doi.org/10.1145/2505515.2527109>.
- Heffernan, K., & Teufel, S. (2018). Identifying problems and solutions in scientific text. *Scientometrics*. <https://doi.org/10.1007/s11192-018-2718-6>.
- Kacem, A., & Mayr, P. (2018). Analysis of search stratagem utilisation. *Scientometrics*. <https://doi.org/10.1007/s11192-018-2821-8>.
- Karimi, S., Moraes, L., Das, A., Shakery, A., & Verma, R. (2018). Citance-based retrieval and summarization using IR and machine learning. *Scientometrics*. <https://doi.org/10.1007/s11192-018-2785-8>.
- Ma, S., Xu, J., & Zhang, C. (2018). Automatic identification of cited text spans: a multi-classifier approach over imbalanced dataset. *Scientometrics*. <https://doi.org/10.1007/s11192-018-2754-2>.
- Mayr, P., Chandrasekaran, M. K., & Jaidka, K. (Eds.) (2017a). *BIRNDL'17: Proceedings of 2nd joint workshop on bibliometric-enhanced information retrieval and natural language processing for digital libraries—Co-located with the 40th international ACM SIGIR conference on research and development in information retrieval (SIGIR'17)*, vol. 1888. *CEUR workshop proceedings*. <http://ceur-ws.org/Vol-1888>.
- Mayr, P., Frommholz, I., & Cabanac, G. (Eds.) (2017b). *BIR'17: Proceedings of the 5th international workshop on bibliometric-enhanced information retrieval—Collocated with the 39th European conference on information retrieval (ECIR'17)*, vol. 1823. *CEUR workshop proceedings*. <http://ceur-ws.org/Vol-1823>.
- Mayr, P., Frommholz, I., Cabanac, G., Chandrasekaran, M. K., Jaidka, K., Kan, M. Y., et al. (forthcoming). Introduction to the special issue on bibliometric-enhanced information retrieval and natural language processing for digital libraries (BIRNDL). *International Journal on Digital Libraries*. <https://doi.org/10.1007/s00799-017-0230-x>.
- Mayr, P., & Scharnhorst, A. (2015). Combining bibliometrics and information retrieval: Preface. *Scientometrics*, 102(3), 2191–2192. <https://doi.org/10.1007/s11192-015-1529-2>.
- Rexha, A., Kröll, M., Ziak, H., & Kern, R. (2018). Authorship identification of documents with high content similarity. *Scientometrics*, 115(1), 223–237. <https://doi.org/10.1007/s11192-018-2661-6>.
- Tien, N. M., & Labbé, C. (2018). Detecting automatically generated sentences with grammatical structure similarity. *Scientometrics*. <https://doi.org/10.1007/s11192-018-2789-4>.
- Yang, S., Han, R., Wolfram, D., & Zhao, Y. (2016). Visualizing the intellectual structure of information science (2006–2015): Introducing author keyword coupling analysis. *Journal of Informetrics*, 10(1), 132–150. <https://doi.org/10.1016/j.joi.2015.12.003>.