

Bibliometric-enhanced Information Retrieval

Ingo Frommholz, Philipp Mayr, Guillaume Cabanac, Suzan Verberne

▶ **To cite this version:**

Ingo Frommholz, Philipp Mayr, Guillaume Cabanac, Suzan Verberne. Bibliometric-enhanced Information Retrieval: 12th International BIR Workshop (BIR 2022). 44th European Conference on Information Retrieval (ECIR 2022), Apr 2022, Stavanger, Norway. pp.561-565, 10.1007/978-3-030-99739-7_70 . hal-03652739

HAL Id: hal-03652739

https://ut3-toulouseinp.hal.science/hal-03652739

Submitted on 27 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Bibliometric-enhanced Information Retrieval: 12th International BIR Workshop (BIR 2022)

Ingo Frommholz[1], Philipp Mayr[2], Guillaume Cabanac[3], and Suzan Verberne[4]

[1] School of Mathematics and Computer Science, University of Wolverhampton, UK,
ifrommholz@acm.org
[2] GESIS – Leibniz-Institute for the Social Sciences, Cologne, Germany,
philipp.mayr@gesis.org
[3] University of Toulouse, Computer Science Department, IRIT UMR 5505, France
guillaume.cabanac@univ-tlse3.fr
[4] Leiden University, LIACS, Leiden, the Netherlands
s.verberne@liacs.leidenuniv.nl

Abstract. The 12th iteration of the Bibliometric-enhanced Information Retrieval (BIR) workshop series is a full-day ECIR 2022 workshop. BIR tackles issues related to, for instance, academic search and recommendation, at the intersection of Information Retrieval, Natural Language Processing, and Bibliometrics. As an interdisciplinary scientific event, BIR brings together researchers and practitioners from the Scientometrics/Bibliometrics community on the one hand and the Information Retrieval community on the other hand. BIR is an ever-growing topic investigated by both academia and the industry.

Keywords: Academic Search, Information Retrieval, Digital Libraries, Bibliometrics, Scientometrics

1 Motivation and Relevance to ECIR

The aim of the BIR workshop series is to bring together researchers and practitioners from Scientometrics/Bibliometrics as well as Information Retrieval (IR). Scientometrics is a sub-field of Bibliometrics which, like IR, is, in turn, a sub-field of Information Science. Bibliometrics and Scientometrics are concerned with all quantitative aspects of information and academic literature [6], which naturally make them interesting for IR research, in particular when it comes to academic search and recommendation. In the 1960s, Salton was already striving to enhance IR by including clues inferred from bibliographic citations [7]. In the course of decades, both disciplines (Bibliometrics and IR) evolved apart from each other over time, leading to the two loosely connected fields we know of today [8].

However, the exploding number of scholarly publications and the need to satisfy scholars' specific information needs led Bibliometric-enhanced IR to receive growing recognition in the IR as well as the Scientometrics communities. Challenges in academic search and recommendation became particularly apparent during the COVID-19 crisis and the need for effective and efficient solutions

for scholarly search for and discovery of high-quality publications on that topic. Bibliometric-enhanced IR tries to provide these solutions to the peculiar needs of scholars to keep on top of the research in their respective fields, utilising the wide range of suitable relevance signals that come with academic scientific publications, such as keywords provided by authors, topics extracted from the full-texts, co-authorship networks, citation networks, bibliometric figures, and various classification schemes of science. Bibliometric-enhanced IR systems must deal with the multifaceted nature of scientific information by searching for or recommending academic papers, patents, venues (i.e., conference proceedings, journals, books, manuals, grey literature), authors, experts (e.g., peer reviewers), references (to be cited to support an argument), and datasets.

To this end, the BIR workshop series was founded in 2014 [5] to tackle these challenges by tightening up the link between IR and Bibliometrics. We strive to bring the 'retrievalists' and 'citationists' [8] active in both academia and industry together. The success of past BIR events, as shown in Table 1, evidences that BIR@ECIR is a much needed interdisciplinary scientific event that attracts researchers and practitioners from IR and Bibliometrics alike.

Table 1. Overview of the BIR workshop series and CEUR proceedings

Year	Conference	Venue		Papers	Proceedings
2014	ECIR	Amsterdam,	NL	6	Vol-1143
2015	ECIR	Vienna,	AT	6	Vol-1344
2016	ECIR	Padua,	IT	8	Vol-1567
2016	JCDL	Newark,	US	10 + 10[a]	Vol-1610
2017	ECIR	Aberdeen,	UK	12	Vol-1823
2017	SIGIR	Tokyo,	JP	11	Vol-1888
2018	ECIR	Grenoble,	FR	9	Vol-2080
2019	ECIR	Cologne,	DE	14	Vol-2345
2019	SIGIR	Paris,	FR	16 + 10[b]	Vol-2414
2020	ECIR	Lisbon (Online),	PT	9	Vol-2591
2021	ECIR	Lucca (Online),	IT	9	Vol-2847

[a] with CL-SciSumm 2016 Shared Task; [b] with CL-SciSumm 2019 Shared Task

2 Workshop goals/objectives

Our vision is to bring together researchers and practitioners from Scientometrics/Bibliometrics on the one hand and IR on the other hand to create better methods and systems for instance for academic search and recommendation. Our view is to expose people from one community to the work of the respective other community and to foster fruitful interaction across communities. Therefore, in the call for papers for the 2022 BIR workshop at ECIR, we will address, but are not limited to, current research issues regarding 3 aspects of the academic search/recommendation process:

1. User needs and behaviour regarding scientific information, such as:

   - Finding relevant papers/authors for a literature review.
   - Filtering high-quality research papers, e.g. in preprint servers.
   - Measuring the degree of plagiarism in a paper.
   - Identifying expert reviewers for a given submission.
   - Flagging predatory conferences and journals.
   - Understanding information-seeking behaviour and HCI in academic search.
2. Mining the scientific literature, such as:
   - Information extraction, text mining and parsing of scholarly literature.
   - Natural language processing (e.g., citation contexts).
   - Discourse modelling and argument mining.
3. Academic search/recommendation systems, such as:
   - Modelling the multifaceted nature of scientific information.
   - Building test collections for reproducible BIR.
   - System support for literature search and recommendation.

3 Target Audience and Dissemination

The target audience of the BIR workshops is researchers and practitioners, junior and senior, from Scientometrics as well as IR and Natural Language Processing (NLP). These could be IR/NLP researchers interested in potential new application areas for their work as well as researchers and practitioners working with bibliometric data and interested in how IR/NLP methods can make use of such data. BIR 2022 will be open for anyone interested in the topic.

The 10th-anniversary edition in 2020 ran online with an audience peaking at 97 online participants [1]. BIR 2021, the 11th edition [3,4], attracted around 57 participants at peak times but a larger number throughout due to participants dropping in and out.

In December 2020, we published our third special issue emerging from the past BIR workshops [2]. More special issues based on BIR workshops are planned.

As a follow-up of the workshop and following the tradition of previous years, the co-chairs will write a report summing up the main themes and discussions to *SIGIR Forum* [4, for instance] and BCS Informer, as a way to advertise our research topics as widely as possible among the IR community. As in the past, we plan to publish our accepted papers open-access in a CEUR Workshop Proceedings volume.

4 Peer Review Process and Workshop Format

Our peer review process will be supported by Easychair. Each submission is assigned to 2 to 3 reviewers, preferably at least one expert in IR and one expert in Bibliometrics or NLP. The programme committee for 2022 will consist of peer reviewers from all participating communities. Accepted papers are either long papers (15-minute talks) or short papers (5-minute talks). Two interactive sessions close the morning and afternoon sessions with posters and demos, allowing attendees to discuss the latest developments in the field and opportunities (e.g.,

shared tasks such as CL-SciSumm). We also invite attendees to demonstrate prototypes during flash presentations (5 minutes).

These interactive sessions serve as ice-breakers, sparking interesting discussions that, in non-pandemic times, usually continue during lunch and the evening social event. The sessions are also an opportunity for our speakers to further discuss their work. BIR has a friendly and open atmosphere where there is an opportunity for participants (including students) to share their ideas and current work and to receive feedback from the community.

5 Organisers and Programme Commitee

- Ingo Frommholz is Reader in Data Science at the University of Wolverhampton, UK.
- Philipp Mayr is a team leader at the GESIS – Leibniz-Institute for the Social Sciences department Knowledge Technologies for the Social Sciences, Germany.
- Guillaume Cabanac is an Associate Professor at the University of Toulouse, France.
- Suzan Verberne is an associate professor at the Leiden Institute of Advanced Computer Science (LIACS) at Leiden University and group leader of Text Mining and Retrieval Leiden.

The list of PC members will be available on the BIR 2022 page.

References

1. Cabanac, G., Frommholz, I., Mayr, P.: Report on the 10th Anniversary Workshop on Bibliometric-enhanced Information Retrieval (BIR 2020). SIGIR Forum **54**(1) (2020). doi:10.1145/3451964.3451974
2. Cabanac, G., Frommholz, I., Mayr, P.: Scholarly literature mining with Information Retrieval and Natural Language Processing: Preface. Scientometrics **125**(3), 2835–2840 (2020). doi:10.1007/s11192-020-03763-4
3. Cabanac, G., Frommholz, I., Mayr, P., Verberne, S. (eds.): Proceedings of the 11th International Workshop on Bibliometric-enhanced Information Retrieval (BIR). No. 2847 in CEUR Workshop Proceedings, Aachen (2021), http://ceur-ws.org/Vol-2847/
4. Frommholz, I., Cabanac, G., Mayr, P., Verberne, S.: Report on the 11th Bibliometric-enhanced Information Retrieval Workshop (BIR 2021). SIGIR Forum **55**(1) (2021). doi:10.1145/3476415.3476426
5. Mayr, P., Scharnhorst, A., Larsen, B., Schaer, P., Mutschke, P.: Bibliometric-Enhanced Information Retrieval. In: 36th European Conference on IR Research, ECIR 2014, Amsterdam, The Netherlands, Proceedings. pp. 798–801. Springer (2014). doi:10.1007/978-3-319-06028-6_99
6. Pritchard, A.: Statistical bibliography or bibliometrics? [Documentation notes]. Journal of Documentation **25**(4), 348–349 (1969). doi:10.1108/eb026482
7. Salton, G.: Associative document retrieval techniques using bibliographic information. Journal of the ACM **10**(4), 440–457 (1963). doi:10.1145/321186.321188

8. White, H.D., McCain, K.W.: Visualizing a discipline: An author co-citation analysis of Information Science, 1972–1995. Journal of the American Society for Information Science **49**(4), 327–355 (1998). doi:b57vc7