



Special issue on de-personalisation, diversification, filter bubbles and search

Ralf Bierig¹ · Simon Caton²

Received: 7 September 2019 / Accepted: 11 September 2019 / Published online: 18 September 2019
© Springer Nature B.V. 2019

1 Introduction

Information retrieval and recommender systems have resulted in a personalised web experience with resounding success (Liu et al. 2019; Adomavicius and Tuzhilin 2015). Building on context, location and users' virtual (social) profiles, the web is highly aligned to users' perceived interests and to the interests of similar people to whom a user is digitally connected. While this delivers relevant content, it also polarises informational perspectives and removes serendipity through the development of filter bubbles (or echo chambers). Filter bubbles refer to scenarios where specific ideas, beliefs or data are reinforced through repetition of a closed system that limits the free movement of alternative and competing ideas (Pariser 2011). There is the implication that certain ideas or outcomes dominate due to, and resulting in, a bias concerning how specific input is gathered. Under-addressed in the literature are methods to qualify/quantify filter bubbles and the associated effects they have over time. Given that the dividing line between on and offline is becoming increasingly intertwined (Lindner et al. 2015), filter bubbles are an issue for the digital generation as they make life choices based on the digital content they are exposed to.

Discourse on the effects of biased information experiences and echo chambers can often only be found in the topical sphere of (social media) communication (Karlsen et al. 2017). Filter bubbles, however, are specifically relevant to the information retrieval community. Some interest has been developed in providing searchers with serendipitous and diversified search results highlighting increasing awareness of the benefits of broadened information experiences and better search engine performance in relation to the users' information needs (Clarke et al. 2008; Sakai et al. 2013; Santos et al. 2015). Nevertheless, the IR community would further benefit from a better understanding of the underlying structure of such filtered or distorted information spaces (e.g. within document collections and the web) and how these can be studied and evaluated.

✉ Ralf Bierig
ralf.bierig@mu.ie

Simon Caton
simon.caton@ucd.ie

¹ Department of Computer Science, Maynooth University, Maynooth, Ireland

² School of Computer Science, University College Dublin, Dublin, Ireland

Currently, there is no single source that integrates multidisciplinary research that conceptualises and evaluates the bias that results from continuous filtering and personalisation. Whilst there is work in the individual fields of information retrieval, information science, cognitive systems, computational social science and machine learning, there is a lack of work that integrates these domains. This special issue attempts to highlight the need for multidisciplinary research and to contribute towards the discussion of the effects of biased, polarised or (over)filtered information spaces. Specifically, moving forwards, we note an under-addressed need in the literature for:

1. *Reviews* of pertinent aspects of filter bubbles including understanding and determining the needs and boundaries of (de-)personalisation—multidisciplinary research is hard, and as such standardised definitions of key terms, challenges and structured debate on methodologies and the current state of the art are needed for the communities to work together and tackle these issues.
2. *Formal approaches to represent highly personalised filter bubbles* to facilitate experimental approaches, enable user comprehension, and simulate filter bubbles—current work in IR tends to focus on the “optimality” of results. There is a need to be able to create reproducible studies of filter bubbles to better aid our understanding of the socio-technical challenges they foster. Thus, new theoretical and empirical models are needed for use within as well as outside the IR community.
3. *(Meta)studies* are needed that attempt to qualify, quantify and visualise the divergence of users’ personalised search results and information experiences—to address filter bubbles we need mechanisms to identify areas of “concern” to foster debate and give rise to novel approaches to tackle these issues.
4. *Reproducible experimental methods* enabling researchers to investigate and study filter bubbles and formalise any effects of filtered information experiences—new methods are needed to evaluate approaches that foster diversification and de-personalisation. Specifically, studies that consider multiple users or multiple user profiles (search engines, social media, etc.) and contexts (location, time, tasks, devices, etc.) that shed light on the differences in users’ diverging search results and information experience(s). Similarly, more studies into filter bubbles and discussion on the tangible effects and observations of (de-)personalisation is needed to better understand and evaluate new approaches.
5. *New or extended test collections and corpora* are needed to enable and support new approaches seeking to gain experience using, adapting, merging and/or gathering (test) collections and experimental data sets.

2 Papers in this special issue

We received six submissions in response to our calls from which we present the two papers that were accepted.

In Maxwell et al. (2019), the researchers investigate result diversification as a potential means to mitigate the effects of filter bubbles and bias in online search. The user study described in the paper specifically investigates the effects of a diversifying system for aspectual retrieval as an example of complex search. Results show that aspectual searchers using a diversified system submit more queries and examine fewer documents per query. Searchers using the diversifying system were more successful in identifying relevant documents with a wider range of novel aspects.

The paper by Chakraborty et al. (2019) explores the de-personalisation of news recommendations with a closer look to the three metrics of recency, importance (or popularity) and diversity. The authors specifically review the tension between recency and impact — the tendency that a recent story can often not yet provide enough information about its importance, and vice-versa. The paper further proposes a strategy and a practical implementation to estimate future impact, measured by editorial selection and reader popularity, for de-personalised news recommendation.

Acknowledgements We thank all authors for submitting manuscripts to this special issue and choosing it as an outlet for their research in the area of de-personalisation, diversification, filter bubbles and search. We are also thankful to all the reviewers for their efforts in providing useful feedback to the authors. Our gratitude further extends to the editors-in-chief, for their kind support during the preparation and completion of this special issue.

References

- Adomavicius, G., & Tuzhilin, A. (2015). Context-aware recommender systems. In *Recommender systems handbook* (pp. 191–226). Springer US, Boston, MA. https://doi.org/10.1007/978-1-4899-7637-6_6.
- Chakraborty, A., Ghosh, S., Ganguly, N., & Gummadi, K. P. (2019). Optimizing the recency-relevance-diversity trade-offs in non-personalized news recommendations. *Information Retrieval Journal*, 2, 1–29.
- Clarke, C. L., Kolla, M., Cormack, G. V., Vechtomova, O., Ashkan, A., Büttcher, S., & MacKinnon, I. (2008). Novelty and diversity in information retrieval evaluation. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval-SIGIR '08* (p. 659). ACM Press, New York, New York, USA. <https://doi.org/10.1145/1390334.1390446>, <http://portal.acm.org/citation.cfm?doid=1390334.1390446>.
- Karlsen, R., Steen-Johnsen, K., Wollebæk, D., & Enjolras, B. (2017). Echo chamber and trench warfare dynamics in online debates. *European Journal of Communication*, 32(3), 257–273.
- Lindner, A., Hall, M., Niemeyer, C., & Caton, S. (2015). BeWell: A sentiment aggregator for proactive community management work-in. In *CHI'15* (pp. 1055–1060). <https://doi.org/10.1145/2702613.2732787>.
- Liu, J., Liu, C., & Belkin, N. J. (2019). Personalization in text information retrieval: A survey. *Journal of the Association for Information Science and Technology*, 5, asi24234. <https://doi.org/10.1002/asi.24234>.
- Maxwell, D., Azzopardi, L., & Moshfeghi, Y. (2019). The impact of result diversification on search behaviour and performance. *Information Retrieval Journal*, 5, 1–25. <https://doi.org/10.1007/s10791-019-09353-0>.
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. London: Penguin Group.
- Sakai, T., Kando, N., Macdonald, C., & Soboroff, I. (2013). Introduction to the special issue on search intents and diversification. *Information Retrieval*, 16(4), 427–428. <https://doi.org/10.1007/s10791-013-9223-6>.
- Santos, R. L. T., Macdonald, C., & Ounis, I. (2015). Search result diversification. *Foundations and Trends® in Information Retrieval*, 9(1), 1–90. <https://doi.org/10.1561/15000000040>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.