

Workshop on Barriers to Interactive IR Resources Re-use (BIARRR 2019)

Toine Bogers

Aalborg University Copenhagen
Denmark
toine@hum.aau.dk

Maria Gäde

Humboldt-Universität zu Berlin
Germany
maria.gaede@ibi.hu-berlin.de

Vivien Petras

Humboldt-Universität zu Berlin
Germany
vivien.petras@ibi.hu-berlin.de

Samuel Dodson

University of British Columbia
Canada
dodsons@mail.ubc.ca

Mark Hall

Martin-Luther-Universität
Halle-Wittenberg
Germany
mark.hall@informatik.uni-halle.de

Nils Pharo

Oslo Metropolitan University
Norway
nils.pharo@oslomet.no

Luanne Freund

University of British Columbia
Canada
luanne.freund@ubc.ca

Marijn Koolen

Royal Netherlands Academy of Arts
and Sciences
Netherlands
marijn.koolen@di.huc.knaw.nl

Mette Skov

Aalborg University
Denmark
skov@hum.aau.dk

CCS CONCEPTS

• **Information systems** → **Users and interactive retrieval**; •
General and reference → *Empirical studies; Evaluation*;

KEYWORDS

IIR, re-use, secondary use, research design

ACM Reference Format:

Toine Bogers, Samuel Dodson, Luanne Freund, Maria Gäde, Mark Hall, Marijn Koolen, Vivien Petras, Nils Pharo, and Mette Skov. 2019. Workshop on Barriers to Interactive IR Resources Re-use (BIARRR 2019). In *2019 Conference on Human Information Interaction and Retrieval (CHIIR '19), March 10–14, 2019, Glasgow, United Kingdom*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3295750.3298965>

1 DESCRIPTION

What would be the equivalent of a re-usable TREC test collection for the interactive information retrieval (IIR) community? The goal of the BIARRR 2019 workshop is to answer this question by continuing existing community-driven efforts to design and implement a platform for the collection, organization, maintenance, and sharing of resources for IIR experimentation. These efforts developed out of discussions at the CHIIR 2017 workshop on Supporting Complex Search Tasks (SCST 2017) [1] and were expanded upon during the BIARRR workshop at CHIIR 2018 [2, 3].

The IR community has a strong tradition of making research data of system-based experimentation available for re-use, as exemplified by the development of test collections and shared tasks in large-scale initiatives such as TREC, CLEF, NTCIR, and FIRE. These efforts

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHIIR '19, March 10–14, 2019, Glasgow, United Kingdom

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6025-8/19/03.

<https://doi.org/10.1145/3295750.3298965>

have had significant benefits for the IR community, in particular enabling the re-use of aspects of the test collections or shared tasks in other IR research. Comparable efforts have been undertaken to bring this paradigm to IIR research, such as the TREC Interactive and Session tracks, the INEX Interactive track, and the Interactive Social Book Search track. The high degree of variation between IIR studies, however, has meant that none of these have achieved similar degrees of standardisation and re-use. It seems that the traditional shared task structure is not successful in inducing re-use in IIR. Instead the equivalent type and level of re-use is more likely to be achieved through increased sharing of research designs, enabling better comparability, more transparent reporting, and greater methodological standards and rigour.

As stated above, IIR research exhibits a large variety of research designs and methods[11]. This methodological diversity and richness needs to be investigated to identify ways of representing these research materials and also to develop an understanding of how and when researchers currently re-use or would like to find and re-use materials. The workshop environment is the most appropriate venue to gather information and input from the IIR community regarding the kind of methods used, their experiences with both re-using materials and making materials available for re-use, and ideas on how to move towards increased sharing and re-use. The organizers already represent a wide range of IIR research perspectives, but to achieve the goal of encouraging re-use a high degree of community input and commitment is needed.

The goal of the 2nd edition of the BIARRR workshop is to provide an interactive forum to address relevant questions around re-use of IIR materials. At the BIARRR 2018 workshop a number of high-level focus areas were identified [3]. To make further progress on these issues, the focus of the 2019 edition will be on the following five topics:

- **Terminology** The terms used in IIR studies, their definitions, and origins.

- **Research design** The overall research design structures and patterns employed and their potential for re-use.
- **Methodology** General and specific methodologies employed in IIR studies, their origins, and their re-use.
- **Resources** Existing resources re-used as part of IIR studies, how to find them, and issues with re-using them.
- **Reporting** What and how to document aspects of IIR studies to maximise the potential for re-use.

BIIRRR 2019 will provide a highly interactive venue to address these topics, combining short paper presentations with breakout groups. For the presentations we invite short contributions in which authors analyze one of their previously published IIR studies in the context of the five topics listed above. The contributions will be presented in themed sessions in the morning to introduce ideas and potential pitfalls for re-use. The afternoon will then pick up these ideas and discuss them in more depth in breakout groups based around the five themes. The workshop will conclude with a reporting and discussion session as well as planning future activities. This interactive format has proven to be very fruitful in the past at the SCST 2017 and BIIRRR 2018 workshops, acting as the springboard for publications, tools, and funding proposals.

2 RELATED EFFORTS

There have been several successful gatherings directed toward addressing the need for considering how to collect, organize, maintain, and share research resources for conducting IIR experiments. IIR campaigns on this topic include the TREC Interactive Track (1997–2002) [12], the INEX Interactive Track (2004–2010) [13, 14], the Cultural Heritage in CLEF (CHiC) Interactive Task (2013) [15], and the interactive Social Book Search (iSBS) task (2014–2016) [7–9] which provided great insight into the challenges and opportunities for long-term, re-usable IIR research materials.

While these demonstrate the ongoing interest in standardising the evaluation of IIR studies [11], they also show that establishing and maintaining a collaborative platform for the re-use of IIR research instruments is still an open issue. This is due to the complexity of IIR studies, which require a combination of system- and user-centered evaluation approaches [10]. In addition to the tasks and document collections that are needed and provided in most shared tasks, participants, search contexts, tasks, processes, systems, data sets, and evaluation measures all need to be modeled to enable re-use for IIR studies. Building on this work, a subset of this workshop’s organizers are currently in the process of analyzing past IIR studies (published at IliX and CHIIR) for re-use, from which an initial analysis has been submitted as a short paper to CHIIR 2019.

There have been efforts to collect and make available some of these IIR research components. The Repository of Assigned Search Tasks (RepAST)¹ collects, analyzes, and shares search tasks taken from publications of IIR studies. RepAST contains bibliographic data and abstracts from approximately 750 published papers, as well as a list of author-identified search task types (e.g., complex, simple, subject, known-item, factual), and the full text of any assigned search tasks reported in the papers [5]. As such, RepAST serves as a library of tasks. Members of the IIR community are encouraged to

compare and contrast task descriptions as well as reuse the tasks in the collection. While RepAST is valuable, it has been underutilized to date, likely due to a lack of awareness.

Issues related to re-use have also been discussed at various workshops, including the Supporting Complex Search Tasks (SCST) workshops in 2015 and 2017 [1, 6], which were organized based on the experiences of running the iSBS shared task. In particular, the discussions at the popular SCST 2017 workshop (co-located with CHIIR 2017) identified a strong desire within the IIR community to address the issues around re-use. This led to the BIIRRR 2018 workshop at CHIIR 2018, which focused exclusively on the re-use issue and was very productive, resulting in the publication of a research paper [3], as well as a grant proposal, and spawned several informal follow-up meetings. It also served as a starting point for a concrete, community-driven effort focused on the challenges and opportunities for designing and implementing a platform for the collection, organization, maintenance, and sharing of resources for IIR experimentation. The BIIRRR 2019 workshop aims to build upon the organisers’ experience in running highly interactive workshops that produce concrete outcomes for the IIR community—at least one of this workshop’s organizers was involved in each of the previous workshops. The goal is to provide a forum for exchanging experiences with undertaking and documenting IIR studies and then discussing how these can be structured in a platform for IIR re-use.

3 CALL FOR PAPERS

In order to incorporate the community’s experience we invite original contributions in the form of *experience papers* that detail methodological and re-use aspects of previously published or in-press IIR studies. Rather than focus on research questions and results, *experience papers* should focus on the following aspects of IIR studies, which are generally under-reported in scientific publications:

- **Terminology** What terminology did you use to describe the different components of the study? Why did you choose this terminology? How did you develop this terminology?
- **Methodology** What overarching and specific methodologies did you employ in the study? How did you decide which methodologies to employ? Examples of overarching methodologies include qualitative/quantitative/mixed methods, theory/practice/design, distant/close reading, big data/small data. Specific methods include, for instance, log studies, eye tracking, A/B testing, and simulated work tasks.
- **Research designs** What research design(s) did you use? Which (aspects) of these have the potential to be re-used? To capture the variation within the IIR field, we use the broad definition of research designs from Cheek [4]: “*the way in which a research idea is transformed into a research project or plan that can then be carried out in practice by a researcher or research team*”.
- **Re-use** What previously created materials did you re-use? This can cover all aspects, such as research designs, software, interfaces, data, scales, and specific survey questions. How did you decide what to re-use? How did you discover the materials that you re-used? Which problems did you encounter searching for them?

¹<https://ils.unc.edu/searchtasks/search.php>

- **Documentation** What aspects of your study could be re-used and how have you documented and represented them to enable re-use? What aspects were fully documented in your publication? What aspects do you feel should be documented outside the main publication?

We invite contributions from both early career and established academics to develop a well-rounded picture of the challenges and rewards of sharing and re-using aspects of IIR studies at different stages in the academic career path.

3.1 Important dates

- **First call for participation:** November 19, 2018
- **Submission deadline:** January 18, 2019
- **Notification of acceptance:** February 4, 2019
- **Camera-ready deadline:** February 18, 2019
- **Workshop at CHIIR 2019:** March 14, 2019

4 WORKSHOP ACTIVITIES

Like the 2018 edition, BIIRRR 2019 will be a highly interactive, full-day workshop, which will combine accepted presentations, discussion lead-ins, and break-out discussions. We will start the day with a full round of introductions of all participants, asking them to identify their interest in the workshop. The morning program will contain up to 10 presentations of accepted *experience papers*, divided over two paper sessions, organized to elicit more interaction and discussion. Each paper session will contain up to five papers and will follow the same format: each paper will be presented (10 minutes), followed by brief clarifying questions (3 minutes) about it while the next presenter sets up. During the last 20 minutes of the session, the session moderator will introduce and moderate a panel discussion among the presenters to probe the experiences and interplay of the presented work and positions. Each presenter will be asked in advance to prepare at least one question about the other papers. The afternoon is organized around two break-out sessions dedicated to discussion of the five workshop themes introduced in Section 3. The workshop will close with a concrete mission statement and a clear plan for future work in this area.

The workshop targets a broad audience of researchers at any level of their academic careers, working in IIR as well as related fields dealing with heterogeneous research designs. In order to achieve this, we plan to directly target experienced researchers in the field and encourage them to submit their own work and also get their doctoral and post-doctoral researchers to contribute. We expect around 10-15 paper submissions, and have room to accept 8-10. To enable an interactive workshop where everyone is involved, we aim to have 25-30 participants with breakout groups of 6-8 people—thereby broadening the scope compared to the 2018 edition of the BIIRRR workshop.

4.1 Tentative workshop program

The workshop is planned as a **full-day** workshop with a tentative program shown below:

- **Introduction**
 - Short introductions by participants (all participants)
 - Introduction of workshop purpose, goals and planned activities (workshop leads)

- Presentation of analysis of IIR research at previous IiX and CHIIR conferences (workshop leads)

- **Paper session 1**

- Presentation of 4-5 experience papers with 10 mins for presentation and 3 mins for clarification questions
- Panel discussion between the paper presenters and the audience about the sessions' experience papers

- **Coffee break**

- **Paper session 2**

- Presentation of 4-5 experience papers with 10 mins for presentation and 3 mins for clarification questions
- Panel discussion between the paper presenters and the audience about the sessions' experience papers

- **Lunch break**

- **Break-out session 1**

- Introduction to break-out round 1
- Break-out discussions round centered around the five workshop themes
- Feedback round: feedback from each break-out group (all participants)

- **Coffee break**

- **Break-out session 2**

- Introduction to break-out round 2
- Break-out discussions round centered around the five workshop themes
- Feedback round: feedback from each break-out group (all participants)

- **Closing**

- Summary of insights (workshop leads)
- Discussion & task assignment for next steps (all participants)
- Collaborative composition & publication of closing statement (workshop leads)

The break-out discussions will have smaller groups, organized around the five workshop themes. We will gauge interest in the different themes from the workshop participants before the start of the workshop and this, together with the interests identified in the short introductions by all participants, will be used to assign participants to the break-out discussions. This should allow us to ensure that each break-out group represents different experience levels and also research backgrounds, with the aim of ensuring that the results from the break-out groups' discussions are as representative as possible.

Written comments and minutes for the workshop will be solicited through a shared document, which will be continuously updated throughout the workshop and which will be available from the workshop website. Via this document, Twitter, and other available social media channels, interested IIR researchers may participate remotely if they cannot attend the workshop in person.

5 DESIRED OUTCOMES & CONTINUING ACTIVITIES

The 2018 edition of the BIIRRR workshop defined five working areas to move the discussion on IIR resource re-use forward [3], one of them being a discussion of the aspects of research designs—more broadly the diversity of methodological approaches employed by

the community. BIIRRR 2019 is designed around developing specific requirements and activities to move this area forward.

One of the main outcomes of BIIRRR 2019 will therefore be a set of requirements on *terminology*, *methodology*, and *research designs* that will be developed in the breakout groups, drawing on the experience paper presentations, the results of the IIR re-use survey conducted in the context of BIIRRR 2018, and an ongoing analysis of IIR studies. The focus of the outcome is on producing concrete requirements that can directly feed into the development of the so-called *iRepository*—a long-term repository of research designs and components used in IIR research. The aim is not to produce a definitive, final set of requirements, but an initial set of requirements that form the baseline from which to develop and then further evolve the *iRepository*.

The second outcome will be a set of principles around the themes of *re-use* and *documentation*, which will provide specific guidelines on how the requirements for the three themes listed above are to be integrated into the *iRepository* to enable realistic re-use. The two outcomes will be published in a short report on the workshop website, followed by a more detailed SIGIR Forum publication, to ensure wide-spread dissemination within the IR community and to encourage engagement with the development of the *iRepository*.

The central aspect of these two initial outcomes is that they will enable follow-up events to have a solid foundation from which to continue, rather than having to re-invent the wheel each time. As part of this—and to ensure that the workshop’s momentum is maintained—we propose setting up a public mailing list. This will not only allow the discussions to continue after the workshop completes, it will also enable researchers who could not attend the workshop to engage in the discussion, increasing the likelihood that the workshop achieves its aims.

In addition to these short-term outcomes, a community-wide repository requires a certain amount of discussion and coordination of all its development. To achieve this, we envision a more extensive activity, such as a Dagstuhl-style seminar, which will provide the time to discuss the requirements and solutions in more detail. In parallel to that, smaller hackathon-style events could be organized that focus on developing technical solutions for aspects of the problem. From our experience with the BIIRRR 2018 outcomes, these tend to form organically around existing conferences in the field, and the aim here is to increase the structure and planning of these events to encourage wider participation.

Finally, as part of this workshop, participants will also be invited to commit to participating in or leading some of the *iRepository* development activities. The aim of this is to encourage wider community participation in this process, but also to enable participants to organically form into interest groups with a view towards potential grant applications in their area of interest, as was an unplanned—but very welcome—outcome of the BIIRRR 2018 workshop.

To achieve these outcomes, wide dissemination to the community is a key activity, in particular to maintain momentum directly after the workshop. We propose to provide more dissemination activities on the workshop website directly after its completion. The main

focus of this will initially be a short statement of intent that is produced and discussed at the end of the workshop and published within a few days of the workshop. This statement of intent will not be comprehensive, but indicate the direction of the initiative and providing a point of reference that participants can state their commitment to. This will be followed by the more extensive report and SIGIR Forum contribution described above. Long-term we are aiming for an article in an appropriate journal to articulate our vision, detail the knowledge gained from the BIIRRR 2018 survey, workshop outcomes, results of the IIR study re-use analysis, and the concrete, practical frameworks developed based on the workshops’ results.

REFERENCES

- [1] Nicholas Belkin, Toine Bogers, Jaap Kamps, Diane Kelly, Marijn Koolen, and Emine Yilmaz. 2017. Second Workshop on Supporting Complex Search Tasks. In *Proc CHIIR 2017*. ACM, New York, NY, 433–435.
- [2] Toine Bogers, Maria Gäde, Luanne Freund, Mark Hall, Marijn Koolen, Vivien Petras, and Mette Skov. 2018. Workshop on Barriers to Interactive IR Resources Re-use. In *Proc CHIIR 2018*. ACM, New York, NY, 382–385.
- [3] Toine Bogers, Maria Gäde, Mark Hall, Luanne Freund, Marijn Koolen, Vivien Petras, and Mette Skov. 2018. Report on the Workshop on Barriers to Interactive IR Resources Re-use (BIIRRR 2018). *SIGIR Forum* 52, 1 (Aug. 2018), 119–128.
- [4] Julianne Cheek. 2008. Research design. In *The SAGE Encyclopedia of Qualitative Research Methods*, Lisa M. Given (Ed.). SAGE.
- [5] Luanne Freund and Barbara M. Wildemuth. 2014. Documenting and studying the use of assigned search tasks: RepAST. *Proceedings of the American Society for Information Science and Technology* 51, 1 (2014), 1–4.
- [6] Maria Gäde, Mark M. Hall, Hugo Huurdeman, Jaap Kamps, Marijn Koolen, Mette Skove, Elaine Toms, and David Walsh. 2015. Report on the First Workshop on Supporting Complex Search Tasks. *SIGIR Forum* 49, 1 (June 2015), 50–56.
- [7] Maria Gäde, Mark Michael Hall, Hugo C. Huurdeman, Jaap Kamps, Marijn Koolen, Mette Skov, Toine Bogers, and David Walsh. 2016. Overview of the SBS 2016 Interactive Track. In *Working Notes of the CLEF 2016 Conference (CEUR Workshop Proceedings)*, Krisztian Balog, Linda Cappellato, Nicola Ferro, and Craig Macdonald (Eds.), Vol. 1609. CEUR-WS.org, 1024–1038.
- [8] Maria Gäde, Mark Michael Hall, Hugo C. Huurdeman, Jaap Kamps, Marijn Koolen, Mette Skov, Elaine Toms, and David Walsh. 2015. Overview of the SBS 2015 Interactive Track. In *Working Notes of the CLEF 2015 Conference (CEUR Workshop Proceedings)*, Linda Cappellato, Nicola Ferro, Gareth J. F. Jones, and Eric SanJuan (Eds.), Vol. 1391. CEUR-WS.org.
- [9] Mark Michael Hall, Hugo C. Huurdeman, Marijn Koolen, Mette Skov, and David Walsh. 2014. Overview of the INEX 2014 Interactive Social Book Search Track. In *Working Notes of the CLEF 2014 Conference (CEUR Workshop Proceedings)*, Linda Cappellato, Nicola Ferro, Martin Halvey, and Wessel Kraaij (Eds.), Vol. 1180. CEUR-WS.org, 480–493.
- [10] Diane Kelly. 2009. Methods for Evaluating Interactive Information Retrieval Systems with Users. *Foundations and Trends in Information Retrieval* 3, 1–2 (2009), 1–224.
- [11] Diane Kelly and Cassidy R Sugimoto. 2013. A Systematic Review of Interactive Information Retrieval Evaluation Studies, 1967–2006. *Journal of the American Society for Information Science and Technology* 64, 4 (2013), 745–770.
- [12] Paul Over. 2001. The TREC Interactive Track: An Annotated Bibliography. *Information Processing & Management* 37, 3 (2001), 369–381.
- [13] Nils Pharo, Thomas Beckers, Ragnar Nordlie, and Norbert Fuhr. 2011. Overview of the INEX 2010 Interactive Track. In *INEX '10: Proceedings of the Ninth International Workshop of the Initiative for the Evaluation of XML Retrieval*, Shlomo Geva, Jaap Kamps, Ralf Schenkel, and Andrew Trotman (Eds.). Springer, Berlin, Heidelberg, 227–235.
- [14] Anastasios Tombros, Birger Larsen, and Saadia Malik. 2005. The Interactive Track at INEX 2004. In *INEX '04: Proceedings of the Third International Workshop of the Initiative for the Evaluation of XML Retrieval*, Norbert Fuhr, Mounia Lalmas, Saadia Malik, and Zoltán Szlávik (Eds.). Springer, Berlin, Heidelberg, 410–423.
- [15] Elaine G. Toms and Mark M. Hall. 2013. The CHiC Interactive Task (CHiCi) at CLEF 2013. In *Working Notes of the CLEF 2013 Conference (CEUR Workshop Proceedings)*, Pamela Forner, Roberto Navigli, Dan Tufis, and Nicola Ferro (Eds.), Vol. 1179. CEUR-WS.org.