How we Work, Share, and Re-use at CHIIR

Bogers, Toine; Gäde, Maria; Hall, Mark Michael; Koolen, Marijn; Petras, Vivien; Larsen, Birger

How we Work, Share, and Re-use at CHIIR

Toine Bogers
Department of Computer Science
IT University of Copenhagen
Copenhagen, Denmark
tobo@itu.dk

Maria Gäde
Berlin School of Library and
Information Science
Humboldt-Universität zu Berlin
Berlin, Germany
maria.gaede@ibi.hu-berlin.de

Mark M Hall
School of Computing &
Communications
The Open University
Milton Keynes, United Kingdom

Marijn Koolen
DHLab, KNAW Humanities Cluster
Amsterdam, Netherlands
marijn.koolen@di.huc.knaw.nl

Vivien Petras
Berlin School of Library and
Information Science
Humboldt-Universität zu Berlin
Berlin, Germany
vivien.petras@ibi.hu-berlin.de

Birger Larsen
Department of Communication &
Psychology
Aalborg University Copenhagen
Copenhagen, Denmark
birger@ikp.aau.dk

ABSTRACT

In this paper, we present the results of an initial study of the research, sharing, and re-use practices at the CHIIR conference through a systematic analysis of all CHIIR papers published from 2016 to 2022. We find that CHIIR is a conference predominantly focused on empirical, multi-methods research that over the years has undergone a focusing in terms of the type of research methods that are being used. A modest number of papers re-use existing data and design resources, but infrastructure component re-use is much more rare. Only a fraction of CHIIR papers actually share their own resources, which suggests that there is much to gain in terms of reproducibility of research presented at CHIIR and could potentially be used to support changes in reviewing practices.

CCS CONCEPTS

• **Information systems → Presentation of retrieval results**; **Test collections**; **Users and interactive retrieval**.

KEYWORDS

interactive information retrieval, sharing, re-use

1 INTRODUCTION

In recent years, an increasing attention has been dedicated to the issues of sharing and re-use of scientific research, with mostly focusing on research data. In addition to the formulation and adoption

of the FAIR principles, which promote "Findable, Accessible, Interoperable and Re-usable" data [21], a number of research data repositories have been established [19]. Others have argued, however, that scientific research—and Information Interaction and Retrieval (II&R) research in particular—is a complex process supported by more than just research data. In their 2021 manifesto, Gäde et al. [6] argue that II&R research is supported by three main resource types: (1) research data, (2) research design, and (3) research infrastructure. They proposed eight principles for improving the sharing and re-use of resources in the II&R community.

Little is known, however, about the sharing and re-use practices in II&R for these three resource types. In this paper, we take the *Conference on Human Information Interaction and Retrieval* (CHIIR) conference as our object of study, which is one of the premier publication venues and communities for II&R research. We present the results of a preliminary analysis of seven years of CHIIR publications and address the following research questions in this work:

RQ1 What are the main research foci of research published at the CHIIR conference?

RQ2 What types of research types and designs are most common at CHIIR?

RQ3 Which types of resources are shared and re-used in CHIIR publications and how often?

In the next section related work on sharing and re-use, as well as analyses of other conferences is highlighted. We then describe our methodology in Section 3 and present our results and conclusions in Sections 4 and 5.

2 RELATED WORK

Analyses of research conferences and communities. Research on communities, practices, publication and citation as well as reviewing processes have revealed general and specific characteristics of different disciplines and relevant conferences. Previous research has leveraged a variety of data sources such as keywords to identify research themes [10], or combinations of subject headings and metadata information as well as scientometric approaches with respect to specific characteristics such as diversity, geographic origin and gender distribution [1, 11–14]. Peer review data from the iConference studies has been analyzed focusing on demographic

and regional influences [5] as well as geographic mobility [8] of researchers. With a specific focus on IIR studies, Petras et al. [17] analyzed a subset of IIiX and CHIIR conference publications to identify methodological components of experimental IIR studies. Most studies argue that meta data alone and especially keywords provided from users are not always meaningful and differ significantly in accuracy.

Sharing & re-use of research resources. While research data access, re-use, and reproducibility have been promoted in the context of open science and data movements for years, there is still a lack of knowledge about the intentions, practices and barriers to re-use [16, 18, 20]. Even less is known about the sharing and re-use patterns of research design and infrastructure components. In the context of the II&R community, the BIIRRR workshop series has aimed at providing an interactive forum for addressing questions around barriers as well as requirements for documentation and re-use of II&R materials, as a step towards the development of community standards [2–4]. Impressions and results from the workshop series resulted in a manifesto on resource re-use in interactive information retrieval [6], which highlights the potential and challenges of research documentation and archiving for future use. However, in addition to the prescriptive manifesto, it is clear that a better descriptive understanding of actual sharing and re-use behavior in the II&R community is necessary as well. To the best of our knowledge, this paper represents the first attempt to analyze the CHIIR community's sharing, re-use, and publication practices through a systematic analysis of the first seven years of CHIIR proceedings.

3 METHODOLOGY

In order to study sharing and re-use behavior in the II&R community and how it intersects with the different research traditions in this community, we chose the CHIIR conference series. CHIIR is a relatively young conference and represents the intersection of several communities. The inaugural CHIIR conference was organized in 2016 in Chapel Hill, NC, and was the result of a successful merger of the *Information Interaction in conteXt* (IIiX) conferences and the *Human Computer Information Retrieval* (HCIR) symposia, which had been organized since 2006 and 2007 respectively. For this short paper, we view the CHIIR conference papers as a first approximation of the research output for this community.

3.1 Data Collection

To undertake our analysis, we downloaded all 355 full, short, perspective, demonstration, and resource papers for the complete CHIIR conference series between 2016-2022. We integrated the bibliographic metadata and full-text of all papers in a Zotero library[1], which was shared between all authors.

3.2 Data Annotation

We annotated the collection of papers in reverse-chronological order by CHIIR proceedings—starting with 2022 and ending with 2016—with each paper annotated by a single annotator. Within each proceedings year, no order was enforced in the annotation process and no steps were undertaken to structure the allocation of papers

[1]https://www.zotero.org/

to annotators. All authors annotated papers for discussion, but four of the authors performed the bulk of the annotation. We created a coding scheme that covered sharing and re-use, research type, research method, and research focus, each of which are described in more detail below. None of the codes were mutually exclusive.

Sharing & re-use. For sharing and re-use, we distinguished between the three types of research resources defined by [6]: data, design, and infrastructure. For annotation purposes, *research design* was defined as "the methods and techniques used to collect and analyse empirical data", *research infrastructure* as the technical infrastructure providing "access to an IR system as well as the application of the data collection techniques", and *research data* as "any data that has been collected, observed, generated or created during or as a results of the research process" [6].

Re-use of one of the three resource types is understood as any use after the initial publication. This may include the re-use of a particular scale (e.g., the User Engagement Scale [15] as part of a post-study questionnaire), the re-use of a particular TREC collection as research data re-use in a lab-based study, or the re-use of an open source IR system [7] as infrastructure re-use.

Sharing is defined as providing access to any resource type to allow future re-use. While Gäde et al. [6] define different levels of sharing and re-use, we did not annotate the level of sharing and reuse in this preliminary study. Instead any example of sharing or reuse is annotated as such (distinguished by resource type), regardless whether it is publishing source code for research infrastructure components on GitHub, making research data available on Zenodo, or even just listing the full-text of a simulated work task in the paper itself.

Research type. Our initial categorization of research types was based on Kelly's definitions of exploratory, descriptive and explanatory research [9]. To this, we added three additional types: theoretical (most common in perspective papers), predictive (i.e., training machine learning models for specific purposes), and resource papers (represented by resource and demo papers). We included Kelly's definitions of the first three research types in our codebook directly; for the other three types, we formulated short definitions before annotation started.

Research method. We used an iterative, open coding approach to arrive at our research method categorization, bootstrapped with a set of common research methods (e.g., questionnaires, user studies, diary studies). Throughout the annotation process, we added additional methods when they were identified in the papers. All annotators were notified when a method was added and the addition was discussed if necessary. At the end of the annotation process, we consolidated the list of research methods into a final set of categories. For this study, we concentrated solely on methods used in the data collection phase. In future work, we intend to extend this to also include data analysis methods.

Research focus. A similar approach was initially taken for annotating the research foci, bootstrapping with common foci (e.g., health, education, interactive information retrieval) and then iteratively adding foci as and when required. In a post-annotation step, these individual foci have been grouped into five larger categories.

These research focus category codes indicate whether a paper focuses on specific document aspects (e.g. genres like news or recipes, or types like text or video), on specific task types or user groups, or on specific aspects of II&R research. The categories are discussed in more detail in Section 4.3.

3.3 Data Cleaning & Analysis

After extracting all the codes for all 355 papers from our shared Zotero library including all paper metadata, we performed basic data cleaning of the dataset. Zotero was unable to identify three papers by DOI, so these were added manually. We also added the CHIIR paper categories (long, short, perspective, demo, and resource) to the dataset. After the consolidation phase of our coding scheme, we also performed this consolidation on the codes in our annotated dataset. We formulated whitelists of accepted codes for research method, type, focus and sharing/re-use and filtered our dataset using these lists. All analyses in the rest of the paper were performed on this filtered dataset[2]. All analyses were performed using a combination of Python and R.

4 RESULTS

We now present the results of the initial analysis of the resulting dataset. At this point, the results are descriptive of the state of research in II&R, but work is ongoing to allow for more in-depth conclusions to be reached.

4.1 Research Type

When annotating the CHIIR papers for research type, we used well-established definitions of empirical research design types by Kelly et al. [9] to guide us, such as exploratory, descriptive and explanatory research. A post-hoc review of our research design type annotations indicated a great amount of variety in how these codes were applied by the four annotators with boundaries between descriptive and explanatory research, between exploratory and descriptive research, and between predictive and explanatory research that were considerably more blurred than expected beforehand.

As a result, we did not feel a detailed analysis of these four empirical research types—exploratory, descriptive, explanatory, and predictive—would be defensible. Instead, we aggregated the four categories into a single research type called 'Empirical' for the analysis. Figure 1 shows the distribution of research types for the complete data-set. It shows that CHIIR predominantly—and perhaps unsurprisingly given both its nature and our aggregation—covers *empirical* research ($n = 299$), followed by a smaller number of *resource* papers ($n = 44$, encompassing both demo and resource papers) and *theoretical* papers ($n = 32$).

Figure 2a shows how the share of these research types has changed over the years. Resource papers saw a big increase from 2018 to 2019, even as the number of accepted publications stayed the same across those years. The relatively recent introduction of perspective papers as a dedicated submission category for more theoretical papers is reflect in the increase in theoretical papers in more recent years. This seems to have come mostly at the expense of resource papers. Another possible explanation for the increase in theoretical

---

[2]We make our dataset and source code for the analysis scripts available at https://github.com/marijnkoolen/CHIIR-2023-Sharing-Citing/.
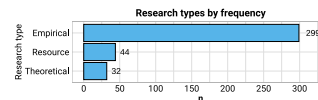

Figure 1: Ordered bar chart of the frequency of different research types. Papers can adopt multiple research types, so total count exceeds the total number of papers ($N = 355$).

papers in 2020 and especially 2021 could be the COVID pandemic, which made it harder to conduct face-to-face empirical research, but which seems to have bounced back in 2022.

4.2 Research Method

The dominance of empirical research is also reflected in in the data collection methods used as shown in Figure 3a. Interestingly, most empirical work used a mixed-methods approach, with a median of 2 ($M = 2.1$) methods per empirical research papers and a maximum of 6 codes, while resource papers were mostly restricted to a median of 1 ($M = 1.2$) method, as were theoretical papers ($M = 1.0$). Figure 3b shows the distribution of the number of data collection methods combined per paper. Questionnaires are the most commonly used method, in part as they are easily included as part of a controlled experiment, user study, usability test or other multi-methods design.

Figure 2b shows how the popularity of different data collection methods has waxed and waned over the years. We focused on the top-seven methods (as shown in Figure 3a) and grouped all other methods together under 'Other'. While there is an expected degree of variety in relative popularity from year to year—since 2018 controlled experiments, interviews and questionnaires have gained in relative popularity whereas user studies and all other methods (as represented by 'Other') have slowly declined in popularity. This may be an effect of the youth of the CHIIR conference and a focusing of the research community.

4.3 Research Focus

The initial analysis of the large set of research focus codes revealed that these coded could be subdivided into five categories:

Application Papers that focus on specific technologies (e.g., digital libraries, conversational agents, recommender systems), or application features (e.g., query suggestion, personalization).

Document Papers that focus on interaction with collections of documents of a certain media type (e.g., data, text, image, audiovisual), genre (e.g., news, recipes) or topic (e.g., aerospace, history).

Research Papers that focus on aspects of II&R research, related to the research process (e.g., evaluation, relevance assessment) or topics (e.g., UX, user engagement, tasks).

Task Papers that focus on specific types of II&R tasks (e.g., exploratory search, collaborative search) the larger work task (e.g., collaboration, design, learning), or the context (e.g., library, archive, museum, theater).

User Papers that focus on specific user groups (e.g., children, elderly people, people with disabilities) or user issues (e.g., distraction, memory, information literacy)
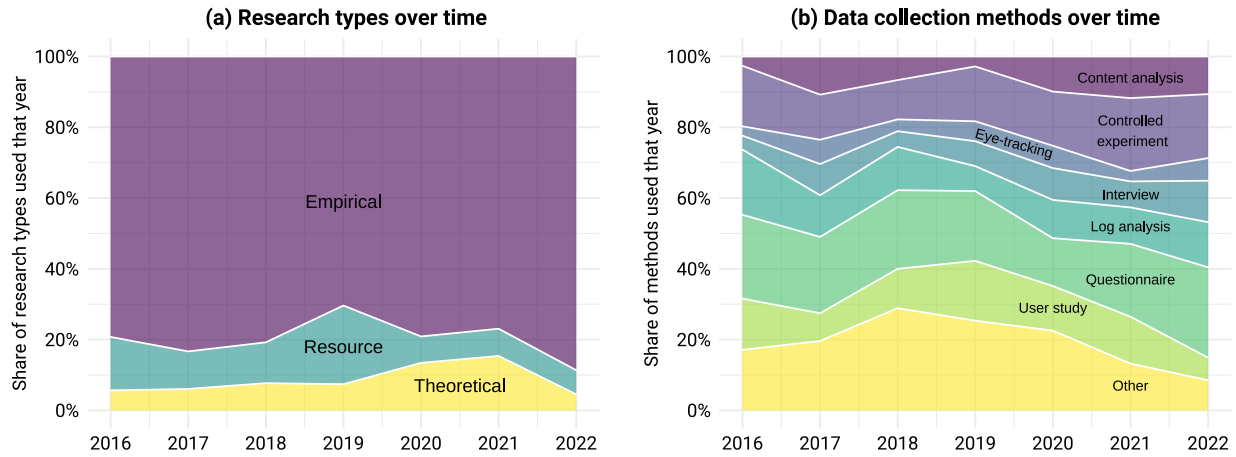
Figure 2: The change in share of used research types (a) and data collection methods (b) over time.
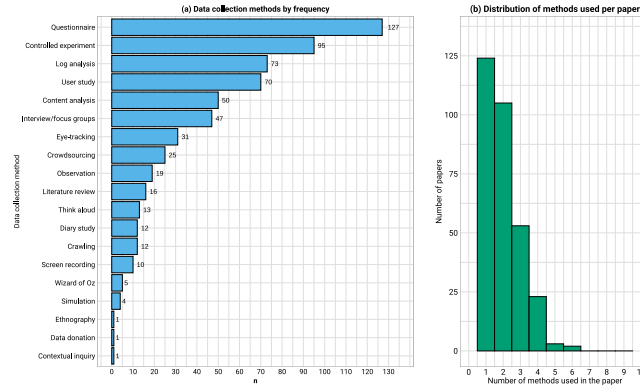

Figure 3: Overview of (a) the frequency of different data collection methods, and (b) distribution of the total number of methods used per paper. Papers can adopt multiple data collection methods, so total count in (a) exceeds the total number of papers (N = 355).

Each paper is annotated with at least one of the five categories, but categories are not mutually exclusive. The median number of codes is 1 (M = 1.6) with a maximum of 5 codes. There are some differences in the numbers of codes assigned to papers from different years. The median (mean) is 2 (M = 1.9) for papers in 2016 and 2017, but drops to 1 ($M \in [1.2, 1.4]$) for papers since 2018. The rising number of codes could be explained by a learning effect resulting from coding the CHIIR proceedings in reverse chronologically order. However, it could also reflect a focusing of the research community as the CHIIR conference matures.

Figure 4 shows the change in category use over time. We can see that document-related papers are consistently most common. In 2016, task and research papers covered a similar fraction, but the fraction of research-focused papers had declined quite strongly and there is also a reduction in the number of task-focused papers. Again, these declines are potentially indicative of a focusing of the research community.
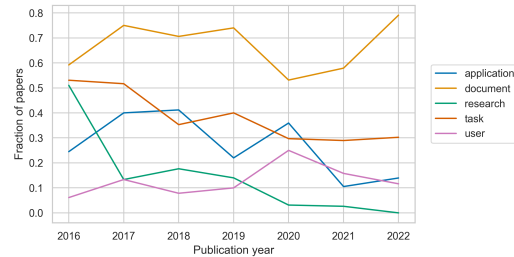

Figure 4: Annual fraction of CHIIR papers focusing on each of the five main domain/topic categories.

4.4 Sharing & Re-use

For sharing and re-use, we distinguished between the three types of research resources defined by [6]: data, design, and infrastructure. Over the seven years of CHIIR, 14.4% of papers re-used existing data,
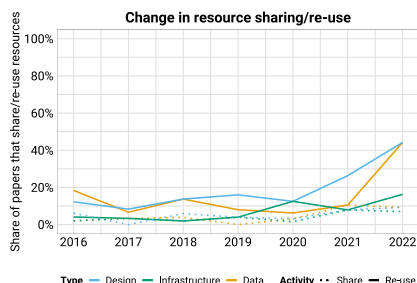
Figure 5: Change in resource sharing and re-use at CHIIR over time as a percentage of all papers ($N = 355$). The dotted lines indicate sharing activity grouped by resource type, while the solid lines indicate re-use activity.

17.7% re-used research designs, and 7.0% re-used infrastructure. At the same time, 4.2% shared (part of) their research data, 4.8% shared elements of their research design, and 3.7% shared at least some infrastructure components. These numbers suggest that a lot of re-used resources either were not explicitly shared through CHIIR papers or originated outside the CHIIR community. Of particular interest is the much lower fraction of infrastructure re-use (7.0% vs. 14.4-17.7%) when compared to sharing, where the fraction is roughly in line with the other resource types (3.7% vs. 4.2-4.8%). This indicates that re-using infrastructure is much harder or provides less value, which is in line with anecdotal evidence [7].

Looking at the change over time in Figure 5, we see an interesting rise in documented re-use in since 2020 with an especially large increase in 2022. While this finding does not necessarily indicate that more resources are being re-used, it does indicate more re-use being documented. This is supported by the lack of a corresponding rise in sharing in 2022. Apart from this rise, the degree of resource sharing and re-use appears to have been relatively stable during the first seven years of CHIIR.

5 DISCUSSION & CONCLUSIONS

In this paper, we presented an initial analysis of the research types, methods, foci as well as sharing and re-use behavior within the II&R community as represented by the publications at the CHIIR conference series 2016-2022. There are some limitations to our analysis. Our analysis is based on accepted papers and therefore represents a pre-selected percentage of the actual research in this field. It might be interesting to extend the analysis to all submitted papers. The papers were annotated in reverse-chronological order by conference year, each paper was annotated by a single annotator, and no effort was made to balance paper-annotator assignments. Thus we cannot determine whether there is any significant inter-annotator disagreement or annotation order influence. Similarly, we cannot at this point determine whether there are any effects of the paper-annotator assignments. We are currently undertaking work to correct this, however, we feel that our current annotations represent an important intermediary step that requires publication and the resulting discussion. However, those aspects will be analyzed and validated in future work.

Interestingly, the one area where we have identified inter-annotator disagreements is the research type annotation. In particular, as mentioned above, post-annotation discussions highlighted significant disagreement in how the exploratory, descriptive, explanatory, and predictive research types were annotated, even though these were based on explicit definitions from the literature. There are likely to be multiple factors driving this, including overlaps in the definitions of the different research types as well as a lack of clarity in the authors' descriptions of their research designs. How to address this remains an open question.

Across all the facets we investigated, we see patterns indicative of an increasingly precise definition of the field and the conference audience, as the number of methods and research foci decrease over time. With regards to sharing and re-use we see a relatively stable pattern over time, with an uptick in the reporting of re-use in 2022. Whether this an actual increase in re-use or just an increase in reporting it, remains unclear. A second, clear conclusion is that re-use of research infrastructure is significantly lower than the re-use of data and designs, most likely because the effort required for re-use of infrastructure is significantly higher.

As a next step, a more detailed analysis including information about analysis methods, application technology methods, user groups, topics, tasks and many more is planned. Through this, we aim to identify the most important aspects of II&R research elements, with the potential that these could be used to propose keywords and categories that could in the future be used by authors to annotate and classify their work.

REFERENCES

[1] Christoph Bartneck and Jun Hu. 2009. Scientometric analysis of the CHI proceedings *(CHI '09)*. ACM, New York, NY, USA, 699–708. https://doi.org/10.1145/1518701.1518810

[2] 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 (BIIRRR 2019). In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*. ACM, 389–392.

[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] Toine Bogers, Maria Gäde, Mark Michael Hall, Marijn Koolen, Vivien Petras, and Paul Thomas. 2022. Third Workshop on Building towards Information Interaction and Retrieval Resources Re-Use (BIIRRR 2022). In *ACM SIGIR Conference on Human Information Interaction and Retrieval* (Regensburg, Germany) *(CHIIR '22)*. Association for Computing Machinery, New York, NY, USA, 374–376. https://doi.org/10.1145/3498366.3505838

[5] Toine Bogers and Elke Greifeneder. 2016. The iSchool community: a case study of iconference reviews. In *IConference 2016 Proceedings*. iSchools. https://doi.org/10.9776/16247

[6] Maria Gäde, Marijn Koolen, Mark Hall, Toine Bogers, and Vivien Petras. 2021. A Manifesto on Resource Re-Use in Interactive Information Retrieval. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval (CHIIR '21)*. Association for Computing Machinery, New York, NY, USA, 141–149. https://doi.org/10.1145/3406522.3446056

[7] Mark M. Hall. 2019. To Re-use is to Re-write: Experiences with Re-using IIR Experiment Software. In *Proceedings of the CHIIR 2019 Workshop on Barriers to Interactive IR Resources Re-use co-located with the ACM SIGIR Conference on Human Information Interaction and Retrieval, BIIRRR@CHIIR 2019, Glasgow, UK, March 14, 2019 (CEUR Workshop Proceedings, Vol. 2337)*, Toine Bogers, Samuel Dodson, Maria Gäde, Luanne Freund, Mark M. Hall, Marijn Koolen, Vivien Petras, Nils Pharo, and Mette Skov (Eds.). CEUR-WS.org, 19–23. http://ceur-ws.org/Vol-2337/paper3.pdf

[8] Vera Hillebrand and Elke Greifeneder. 2022. International mobility in Library and Information Science. In *Proceedings of the 15th International Symposium of Information Science (ISI 2017)*. 371–373. https://doi.org/10.18452/1464

[9] Diane Kelly et al. 2009. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends® in Information Retrieval* 3, 1–2 (2009), 1–224. https://doi.org/10.1561/1500000012

[10] Yong Liu, Jorge Goncalves, Denzil Ferreira, Bei Xiao, Simo Hosio, and Vassilis Kostakos. 2014. CHI 1994-2013: mapping two decades of intellectual progress through co-word analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3553–3562. https://doi.org/10.1145/2556288.2556969

[11] Simon Mahony and Yaming Fu. 2021. Encouraging Diversity of Dialogue as Part of the iSchools Agenda. In *Diversity, Divergence, Dialogue (LNCS)*, Katharina Toeppe, Hui Yan, and Samuel Kai Wah Chu (Eds.). Springer International Publishing, Cham, 195–206. https://doi.org/10.1007/978-3-030-71305-8_15

[12] Dana McKay and George Buchanan. 2019. Shaking the Tree: Understanding Historic and Future Representation of Women at OzCHI. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction (OZCHI'19)*. Association for Computing Machinery, New York, NY, USA, 412–417. https://doi.org/10.1145/3369457.3369504

[13] Dana McKay and George Buchanan. 2020. Feed the Tree: Representation of Australia-based Academic Women at HCI Conferences. In *Proceedings of the 32nd Australian Conference on Human-Computer Interaction (OzCHI '20)*. Association for Computing Machinery, New York, NY, USA, 263–269. https://doi.org/10.1145/3441000.3441061

[14] Dana McKay, Huiwen Zhang, and George Buchanan. 2022. Who am I, and who are you, and who are we? A Scientometric Analysis of Gender and Geography in HCI. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3491102.3502106

[15] Heather L O'Brien and Elaine G Toms. 2010. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology* 61, 1 (2010), 50–69. https://doi.org/10.1002/asi.21229

[16] Irene V. Pasquetto, Christine L. Borgman, and Morgan F. Wofford. 2019. Uses and Reuses of Scientific Data: The Data Creators' Advantage. *Harvard Data Science Review* 1, 2 (15 11 2019). https://doi.org/10.1162/99608f92.fc14bf2d https://hdsr.mitpress.mit.edu/pub/jduhd7og.

[17] Vivien Petras, Toine Bogers, and Maria Gaede. 2019. Elements of IIR Studies: A Review of the 2006-2018 IIiX and CHIIR Conferences. In *Proceedings of the CHIIR 2019 Workshop on Barriers to Interactive IR Resources Re-use (BIIRRR 2019)*. CEUR-WS.org, Aachen, Germany, 37–41.

[18] Stephanie van de Sandt, Artemis Lavasa, Sünje Dallmeier-Tiessen, and Vivien Petras. 2019. The Definition of Reuse. *Data Science Journal* 18 (2019), 22.

[19] Paul Vierkant, Heinz Pampel, Robert Ulrich, Frank Scholze, Maxi Kindling, Michael Witt, and Kirsten Elger. 2019. re3data.org Registry of Research Data Repositories. https://doi.org/10.5281/zenodo.3514133

[20] Jillian C Wallis, Elizabeth Rolando, and Christine L Borgman. 2013. If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PloS one* 8, 7 (2013), e67332.

[21] Mark Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gaby Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Olavo Bonino da Silva Santos, Philip Bourne, Jildau Bouwman, Anthony Brookes, Tim Clark, Merce Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris Evelo, Richard Finkers, and Barend Mons. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3 (03 2016). https://doi.org/10.1038/sdata.2016.18