The Role of HCI in Reproducible Science: Understanding, Supporting and Motivating Core Practices

Sebastian S. Feger
LMU Munich and CERN
Munich, Germany and Geneva, Switzerland
sebastian.s.feger@cern.ch

Sünje Dallmeier-Tiessen
CERN
Geneva, Switzerland
sunje.dallmeier-tiessen@cern.ch

Paweł W. Woźniak
Utrecht University
Utrecht, the Netherlands
p.w.wozniak@uu.nl

Albrecht Schmidt
LMU Munich
Munich, Germany
albrecht.schmidt@ifi.lmu.de

ABSTRACT

The reproducibility crisis refers to the inability to reproduce scientific experiments and is one of science's great challenges. Alarming reports and growing public attention are leading to the development of services and tools that aim to support key reproducible practices. In the face of this rapid evolution, we envision the unique opportunity for Human-Computer Interaction to impact scientific practice through the systematic study of requirements and moderating effects of technology on research reproducibility. In this paper, we report on the current state of technological and human factors in reproducible science and present challenges and opportunities for both HCI researchers and practitioners to understand, support and motivate core practices.

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland Uk

ACM ISBN 978-1-4503-5971-9/19/05.
https://doi.org/10.1145/3290607.3312905

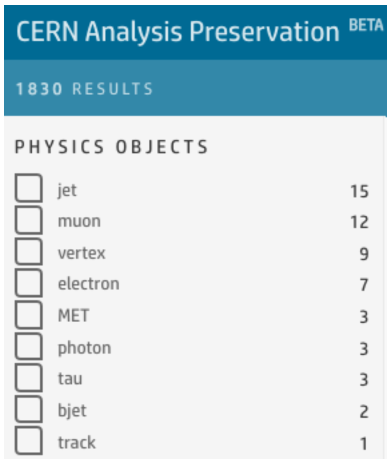

Figure 1: The design of tailored preservation technology needs to reflect search and reuse needs. Here, search facets of CAP allow to access analyses relevant to researchers in selected High Energy Physics experiments.

[1]CERN Analysis Preservation is publicly available: *github.com/cernanalysispreservation*

KEYWORDS

Reproducible Science; Role of HCI; Requirements; Research Challenges; Motivating Practices.

INTRODUCTION

Reproducibility should be a cornerstone of science that enables reuse and independent research verification. Yet, the inability to reproduce research has been reported in numerous scientific fields, from Psychology to Physics [1]. *Preserving* and *sharing* research are core reproducible practices that require efforts to describe, prepare and document resources [5]. But there is often no compelling incentive or reward for researchers to undergo those efforts [7, 9]. Thus, extrinsic motivations, including enforcing submission rules and restrictive funding schemes, are discussed and implemented.

Replication and reuse also concern Human-Computer Interaction (HCI). Recent work investigating open science practices of CHI publications found that resources are rarely shared openly [9]. Such reports add to ongoing discussions concerning the role and value of replication in HCI [14, 22]. But HCI has a unique opportunity to impact reproducible practices beyond its own field through the development of technology required to facilitate preservation and circulation of scientific data [19, 23]. In our research in the context of High Energy Physics (HEP) [11, 12], we learned about requirements and opportunities designing technology for research reproducibility. Future, coordinated efforts are needed to build a general framework that goes beyond the study of reproducible research practices in very specific scientific domains. As we depict in this paper, HCI's methods provide a unique opportunity to systematically study needs and opportunities in science in general.

HCI'S ROLE IN REPRODUCIBLE RESEARCH

Today's availability of online technologies enables institutes, libraries and service providers to develop platforms that support scientists in preserving and sharing their research [19, 23]. *E-Science*, which can broadly be characterized as "the application of computer technology to the undertaking of modern scientific investigation" [2], is expected to enable sharing and reuse [16, 17]. Emerging infrastructure that supports e-science consists primarily of general data repositories and community-tailored services [21]. General repositories are platforms that enable storing of heterogeneous data without imposing a specific structure, thus making them suitable for all types of digital records in a variety of research fields. Community-tailored services are designed to map the researchers' workflows in a specific target domain. CERN Analysis Preservation[1] [6], depicted in Figure 1 and Figure 2, is an example of such a tailored service. The service reflects workflow practices and reuse needs of data analysts working in major HEP experiments. HEP, a strong example of e-Science, is one of science's most data-intensive branches and the context of our research. Tailored services can provide their community with powerful means to preserve, discover and reuse research through domain-tailored language

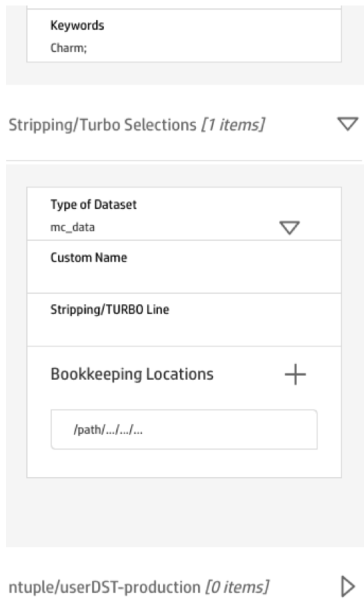

Figure 2: Analysis templates on CAP are tailored to workflow practices of High Energy Physics data analysts. The service is publicly available: *github.com/cernanalysispreservation*

[6, 15]. Studying requirements of preservation technology and physicists' perceptions of CAP, we found that technology can play roles far beyond preservation and reuse [11]. In particular, we found that a tailored service should not only provide tailored submission templates, but also design for community-tailored incentives. Based on our research, we present challenges and opportunities for HCI practitioners and researchers to better understand, support and motivate reproducible practices.

The Role of Practitioners

The design and operation of tools that support reproducibility profits from the involvement of HCI practitioners at all stages of the technologies' life cycle. Based on our involvement in the development and support of preservation technology in HEP, we outline practice-oriented challenges:

Mapping Practices. As part of a platform's design, research workflows need to be well understood. Applicable to both platform types, this is particularly important for tailored tools. Service developers need to involve target communities in the design process, to map submission, search and reuse needs.

Lowering Efforts. Given that the effort to document and share research is a main barrier, data description mechanisms must be supportive. Well-designed submission forms, as well as auto-suggest and auto-complete mechanisms that build on knowledge of research workflows are essential.

Ensuring Usability. Services need to be tested with users as crucial part of the design process, to improve their usability and to detect barriers. This has to be a continuous process, as research description templates on tailored services need to be adapted to novelty and creativity in science.

Providing an Interface. Given the necessary close involvement of practitioners with the research community, HCI practitioners are in an ideal position to interface between technology developers and researchers. Main responsibilities should include: promotion of infrastructure developments to the research community; and feedback communication to the development team.

Emerging Research Challenges

The involvement of HCI practitioners in the design and operation of services for science replicability allows to create a most supportive and efficient interaction with technology. Yet, minimizing the effort is not necessarily enough to engage scientists at large [4]. HCI research has a unique opportunity to impact reproducible practices through the systematic study of requirements and incentive structures:

Understanding Requirements. A pillar of HCI research is based on studying requirements of user groups and populations, allowing to create design models and guidelines. In the context of reproducible research, we need to further understand the requirements and connections of scientific communities and individual researchers to preservation, sharing, infrastructure and knowledge lifecycles [16]. What new forms of community interaction do platforms provide? What common requirements apply to diverse forms of research, including computational, qualitative and descriptive research? And how do requirements differ between these different research forms? Jackson and Barbrow [15] underline the value of studying requirements in field-specific investigations, pointing out the "need to supplement or replace generic, tool-centered, and aspirational accounts of cyberinfrastructure development with approaches that start from the individual histories of practice and value in specific scientific fields."

Incentives / Rewards. Research repositories often advertise a potential citation increase, as emphasizing "the potential of data citations can affect researchers' data sharing preferences from private to more open" [13]. But despite the value of implementing and promoting meaningful benefits, Rowhani-Farid et al. [20] report a lack of incentives in their systematic literature review in the medical domain. They note that even though "data is the foundation of evidence-based health and medical research, it is paradoxical that there is only one evidence-based incentive to promote data sharing." They refer to open science badges [18] and conclude that "more well-designed studies are needed in order to increase the currently low rates of data sharing."

A fundamental understanding of researchers' needs enables description and implementation of new incentive structures. In our research in HEP, we found that technology can create meaningful incentives that profit contributing scientists in their work [11]. Our first interview study points to *secondary usage forms* of preservation technology that can support coping with uncertainty and stimulating useful collaboration. As our research shows, tailored systems can particularly profit from a systematic study of incentives. They are adapted to a particular scientific domain and require a thorough design process in order to provide tailored submission and reuse mechanisms. Extending the research and design process to study and implement also tailored incentives and rewards seems desirable. However, not all institutions and (specialized) research domains can afford implementation and support of highly tailored systems. Thus, as stressed in our paper, future research needs to study more general frameworks for incentives that can be applied across different research forms.

Motivational Design. In recent years, Gamification, "the use of game design elements in non-game contexts" [8] has become a strong example of motivational design in HCI. Badges, one of the most common game design elements, have shown to encourage research data openness in the Psychological Science journal [18]. And also ACM announced the introduction of an even larger set of badges that aim to incentivize reproducible practices [3]. Yet, we have limited knowledge about needs and constraints of gamification in highly skilled scientific environments [10]. Our research on gamified prototypes of a preservation service in HEP shows that gamification can provide motivation if scientific practices are reflected in the design [12]. We contrasted two prototypes in a mixed-method study: While one made use of most common game design elements (including points and leaderboards), the other used a more informative language. Both were rated persuasive and suitable by the experimental physicists that participated in our research. They highlighted how game mechanisms can provide motivation through a fair representation of contributions and best practice efforts. We stressed that future research needs to further study gamification in production preservation systems to learn more about the use of motivational and persuasive design techniques in preservation technology.

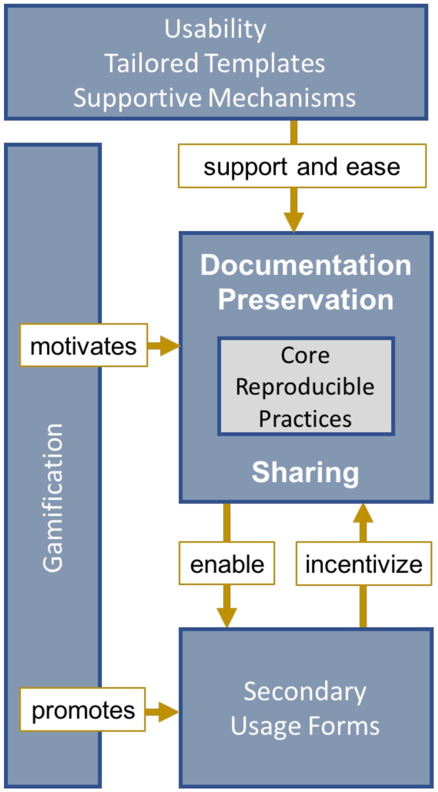

Figure 3: HCI has a unique opportunity to support and motivate reproducible practices through a fundamental understanding of practices around preservation and sharing. Non-conventional design techniques can both create motivation and reshape researchers' perception of the uses of preservation technology.

DISCUSSION & FUTURE WORK

As depicted in Figure 3, we envision an active role of HCI in understanding, supporting and motivating reproducible practices. Taking the example of data analysis in natural science, we described starting points for HCI practitioners and researchers. While it is important to note that our findings result from research in HEP, we argue that the general nature of technology's secondary uses is likely to profit scientific fields beyond experimental physics. We expect HCI to contribute fundamental, empirical knowledge about requirements for technology that supports reproducible research. Research is further expected to describe opportunities and constraints of novel incentive structures and persuasive design techniques in highly-skilled scientific environments. Our vision places further emphasis on the importance of HCI practitioners in implementing and supporting the technological developments.

Taking an active position in reproducible science, we postulate that HCI can further its understanding of the role and value of openness and replication in its own field. In their paper "Is replication important for HCI?", Greiffenhagen and Reeves [14] ask to better understand *aims* and *motivations* of replication in HCI. The authors state that "in order to focus the discussion of replication in HCI, it would be very helpful if one could gather more examples from different disciplines, from biology to physics, to see whether and how replications are valued in these." As HCI researchers and practitioners study and support reproducible efforts in a variety of scientific domains, we will inevitably learn about the value of replication and reuse in numerous fields. In this process, we will also need to learn about varying needs, states and constraints of replication in very different areas of HCI.

ACKNOWLEDGMENTS

This work has been sponsored by the Wolfgang Gentner Programme of the German Federal Ministry of Education and Research.

REFERENCES

[1] Monya Baker. 2016. 1,500 scientists lift the lid on reproducibility. *Nature* 533, 7604 (2016), 452–454. https://doi.org/10.1038/533452a

[2] Shannon Bohle. 2013. What is e-science and how should it be managed. *Nature, Spektrum der Wissenschaft (Scientific American), http://www. scilogs. com/scientific_and_medicallib raries/what-is-e-science-and-how-should-it-be-managed* (2013).

[3] Ronald F Boisvert. 2016. Incentivizing reproducibility. *Commun. ACM* 59, 10 (2016), 5–5.

[4] Christine L. Borgman. 2006. What can studies of e-Learning teach us about collaboration in e-Research? Some findings from digital library studies. *Computer Supported Cooperative Work* 15, 4 (2006), 359–383. https://doi.org/10.1007/s10606-006-9024-1

[5] Christine L Borgman. 2007. *Scholarship in the digital age: information, infrastructure, and the internet.* MIT Press, Cambridge, MA.

[6] Xiaoli Chen, Sünje Dallmeier-Tiessen, Anxhela Dani, Robin Dasler, Javier Delgado Fernández, Pamfilos Fokianos, Patricia Herterich, and Tibor Šimko. 2016. CERN Analysis Preservation: A Novel Digital Library Service to Enable Reusable and Reproducible Research. In *International Conference on Theory and Practice of Digital Libraries.* Springer, 347–356.

[7] Open Science Collaboration. 2012. An Open, Large-Scale, Collaborative Effort to Estimate the Reproducibility of Psychological Science. *Perspectives on Psychological Science* 7, 6 (2012), 657–660. https://doi.org/10.1177/1745691612462588

[8] Sebastian Deterding, Rilla Khaled, Lennart E Nacke, and Dan Dixon. 2011. Gamification: Toward a definition. In *CHI 2011 gamification workshop proceedings*, Vol. 12. Vancouver BC, Canada.

[9] Florian Echtler and Maximilian Häußler. 2018. Open Source, Open Science, and the Replication Crisis in HCI. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* (2018), 1–8. https://doi.org/10.1145/3170427.3188395

[10] Sebastian Feger, Sünje Dallmeier-Tiessen, Paweł Woźniak, and Albrecht Schmidt. 2018. Just Not The Usual Workplace: Meaningful Gamification in Science. *Mensch und Computer 2018-Workshopband* (2018).

[11] Sebastian S. Feger, Sünje Dallmeier-Tiessen, Albrecht Schmidt, and Paweł W. Woźniak. 2019. Designing for Reproducibility: A Qualitative Study of Challenges and Opportunities in High Energy Physics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI'19* (2019). https://doi.org/10.1145/3290605.3300685

[12] Sebastian S. Feger, Sünje Dallmeier-Tiessen, Paweł W. Woźniak, and Albrecht Schmidt. 2019. Gamification in Science: A Study of Requirements in the Context of Reproducible Research. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI'19* (2019). https://doi.org/10.1145/3290605.3300690

[13] Kristian Garza, Carole Goble, John Brooke, and Caroline Jay. 2015. Framing the community data system interface. *Proceedings of the 2015 British HCI Conference on - British HCI '15* (2015), 269–270. https://doi.org/10.1145/2783446.2783605

[14] Christian Greiffenhagen and Stuart Reeves. 2013. Is Replication important for HCI? *CEUR Workshop Proceedings* 976 (2013), 8–13.

[15] Steven J Jackson and Sarah Barbrow. 2013. Infrastructure and Vocation: Field, Calling and Computation in Ecology. *Proc. CHI 2013* (2013), 2873–2882. https://doi.org/10.1145/2470654.2481397

[16] Marina Jirotka, Rob Procter, Tom Rodden, and Geoffrey C. Bowker. 2006. Special Issue: Collaboration in e-Research. *Computer Supported Cooperative Work (CSCW)* 15, 4 (01 Aug 2006), 251–255. https://doi.org/10.1007/s10606-006-9028-x

[17] Helena Karasti, Karen S. Baker, and Eija Halkola. 2006. Enriching the notion of data curation in e-Science: Data managing and information infrastructuring in the Long Term Ecological Research (LTER) network. *Computer Supported Cooperative Work* 15, 4 (2006), 321–358. https://doi.org/10.1007/s10606-006-9023-2

[18] Mallory C. Kidwell, Ljiljana B. Lazarević, Erica Baranski, Tom E. Hardwicke, Sarah Piechowski, Lina Sophia Falkenberg, Curtis Kennett, Agnieszka Slowik, Carina Sonnleitner, Chelsey Hess-Holden, Timothy M. Errington, Susann Fiedler, and Brian A. Nosek. 2016. Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLoS Biology* (2016). https://doi.org/10.1371/journal.pbio.1002456

[19] Irene V. Pasquetto, Ashley E. Sands, Peter T. Darch, and Christine L. Borgman. 2016. Open Data in Scientific Settings: From Policy to Practice. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1585–1596. https://doi.org/10.1145/2858036.2858543

[20] Anisa Rowhani-Farid, Michelle Allen, and Adrian G. Barnett. 2017. What incentives increase data sharing in health and medical research? A systematic review. *Research Integrity and Peer Review* 2, 1 (2017), 4. https://doi.org/10.1186/s41073-017-0028-9

[21] 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). https://doi.org/10.1371/journal.pone.0067332

[22] Max L. L. Wilson, Paul Resnick, David Coyle, and Ed H. Chi. 2013. RepliCHI. *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13* (2013), 3159. https://doi.org/10.1145/2468356.2479636

[23] Daniel J Worden. 2017. Emerging Technologies for Data Research: Implications for Bias, Curation, and Reproducible Results. In *Human Capital and Assets in the Networked World*. https://doi.org/doi:10.1108/978-1-78714-827-720171003