EasyChair Preprint

№ 8519

# Including Data Management in Research Culture Increases the Reproducibility of Scientific Results

Christian Riedel, Hendrik Geßner, Anja Seegebrecht, Safial Islam Ayon, Shafayet Hossen Chowdhury, Ralf Engbert and Ulrike Lucke



July 20, 2022

# Including Data Management in Research Culture Increases the Reproducibility of Scientific Results

Christian Riedel [1], Hendrik Geßner [2], Anja Seegebrecht [3,4], Safial Islam Ayon [5], Shafayet Hossen Chowdhury [6], Ralf Engbert [7], and Ulrike Lucke [8]

**Abstract:** Reproducible research results are among the pillars of sustainable science, and considerable progress has been achieved in this direction recently. However, there is much room for improvement across the research communities. Here we analyze the reproducibility of 108 publications from an interdisciplinary Collaborative Research Center on applied mathematics in various scientific fields. Based on a previous reproducibility study in hydrology, we identify the rate of reproducible scientific results and why reproducibility fails. We identify the main problems that hinder reproducible results and relate them to previous interventions targeting the research culture of reproducible scientific findings. Thus, the success of our measures can be estimated, and specific recommendations for future work can be derived. In our study, the number of publications that allow for at least partly reproducible research results increased over time. However, we see an ongoing need for directives and support in research data management among research communities since issues concerning data accessibility and quality limit the reproducibility of scientific results. We argue that our results are representative of other interdisciplinary research areas.

**Keywords:** open data, reproducibility, research data management

[1] University of Potsdam, Institute of Computer Science, An der Bahn 2, 14476 Potsdam, Germany, christian.riedel@uni-potsdam.de, https://orcid.org/0000-0001-5154-4153
[2] University of Potsdam, Centre for Information Technology and Media Management, Am Neuen Palais 10, 14469 Potsdam, Germany, hendrik.gessner@uni-potsdam.de, https://orcid.org/0000-0002-7786-2587
[3] University of Potsdam, Institute of Computer Science, An der Bahn 2, 14476 Potsdam, Germany, anja.seegebrecht@uni-potsdam.de, https://orcid.org/0000-0002-9404-8528
[4] University of Freiburg, Institute of Physics, Hermann-Herder-Str. 3, 79104 Freiburg im Breisgau, anja.seegebrecht@physik.uni-freiburg.de, https://orcid.org/0000-0002-9404-8528
[5] University of Potsdam, Institute of Computer Science, An der Bahn 2, 14476 Potsdam, Germany, safial.islam.ayon@uni-potsdam.de, https://orcid.org/0000-0002-0970-0455
[6] University of Potsdam, Institute of Computer Science, An der Bahn 2, 14476 Potsdam, Germany, md.shafayet.hossen.chowdhury@uni-potsdam.de, https://orcid.org/0000-0001-5536-9807
University of Potsdam, Department of Psychology, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany, ralf.engbert@uni-potsdam.de, https://orcid.org/0000-0002-2909-5811
[8] University of Potsdam, Institute of Computer Science, An der Bahn 2, 14476 Potsdam, Germany, ulrike.lucke@uni-potsdam.de, https://orcid.org/0000-0003-4049-8088

# 1 The need for and measures to increase reproducibility

The reproducibility of research results is one of the cornerstones of scientific quality assurance [De22]. If a published result is reproducible, this lends credibility to the original study. While most research publications in applied mathematics and the sciences are based on research data and software, many scientific communities currently report low reproducibility rates [St19a, Nü18] or mixed results [SSM18, Ob20, CP16]. This concern, sometimes referred to as a *reproducibility crisis*, is well known within the scientific community [Ba16]. However, institutional examinations on reproducibility rates are rarely carried out, particularly regarding research data management (RDM) interventions [Pe17].

The resulting knowledge gap is more remarkable as the measures to increase reproducibility start at the institutional level, including education and training in various stages of the scientific career [WC19] and support and methodological enhancement for dedicated research groups [Gö19]. While such support initially focused on the provision of IT services [WGS16], there is an increasing orientation towards cultural change in RDM, including research interest in this field [Ge21]. This also applies to the Collaborative Research Center (CRC) 1294 – Data Assimilation, which is the subject of this study.

In general, the term *research data* refers to all kinds of data that arise during the planning, implementation, and documentation of scientific projects or are used in such a project [EU19]. This includes measurements, laboratory values, audiovisual information, text corpora, survey data, observation data, methodological test procedures, and questionnaires. In addition, software and simulations - as in the context of our CRC - also represent central results of scientific research and, therefore, are included in this expression. Dedicated skills are required for the systematic handling of research data throughout the whole process of research. Such education and training are also the subjects of current national RDM initiatives[9].

Within this context, the CRC 1294 – Data Assimilation established such support structures. The CRC is funded by the German Research Foundation and involves 15 research projects, where seven projects have a mathematical focus, and eight projects are located in natural sciences. The overall intention of the CRC is to develop advanced numerical data assimilation methods with potential applications arising from other disciplines, mainly within the natural sciences. With the start of the project, the CRC established a framework that supports the handling and exchange of scientific data. This framework includes providing and coordinating IT platforms to facilitate collaborative work, data sharing and archiving, and offering workshops to raise awareness for research data management and the reproducibility of scientific results. The measures go hand in hand with state-of-the-art approaches in research data management in the respective scientific communities.

[9]https://www.mardi4nfdi.de/ and https://nfdixcs.org/ as for our field

By relating the measures to changes in the reproducibility rate, we aim to determine the success of such interventions. To that end, we investigate the reproducibility of 108 scientific results published between 2017 and 2021 by researchers in the CRC. We use this study to investigate why reproducibility fails and give recommendations on institutional measures to increase the reproducibility of scientific results. The structure of this paper is as follows. Section 2 discusses the methodology and results of previous reproducibility studies. On this basis, the study's design and results within the CRC are presented in section 3. We present a detailed analysis of the reasons underlying reproducibility failures and investigate changes in the reproducibility rate over time to relate the observed change to the interventions on research data handling carried out in the CRC. From this correlation, we evaluate the success of these measures. Our study concludes with recommendations (section 4) for further actions transferable to other research areas.

## 2 Analyzing reproducibility in different fields of science

Investigations on the reproducibility of scientific results have shown that the inability to reproduce scientific findings is a commonly observed phenomenon that affects multiple research fields. Such studies typically attempt to repeat the analyses in peer-reviewed publications from a particular research field, which often involves the analysis of research data. Despite different habits in data handling in the respective communities and slightly diverging definitions of research data and reproducibility, reproducibility studies commonly conclude that the unsuccessful replication of scientific findings is a central problem that needs to be addressed [CL15, CP16, SKL18, Ra19, St19a, Ma20, La22].

Since what scientists perceive as their generated research data varies with different methodological approaches, the assessment of reproducibility criteria typically depends on the research field. In computer science, reproducibility describes the ability of researchers to generate the same results as another research group using their data and experimental setup [ACM20]. In investigations covering various scientific fields, we find slightly different definitions that affect the methodology to evaluate reproducible science. For example, authors may allow for a certain margin when reproducing numerical results [SKL18, La22] or require a data availability statement from a journal publication [St19a]. In the study by [CP16], for example, the authors examined journal and conference publications in computer sciences and defined reproducibility as the ability to build and run computer code without investigating the correctness of the obtained results. Therefore, the authors excluded theoretical papers describing algorithmic approaches that are not backed by computer code. The study by [Ra19], on the other hand, examined the reproducibility of such theoretical papers in machine learning research and translated the presented methods into computer code to investigate barriers in the respective algorithms. A paper was considered reproducible when the

code written by the authors could use reasonable standard programming libraries, and the code confirmed a majority of the claims from the paper.

The studies above show that different perceptions of the term *research data* and different methodological approaches affect the assessment of reproducibility, which inherently introduces subjectivity when considering numerical reproducibility rates. For example, the rates of reproducible scientific results can vary dramatically from less than two [St19a] to more than 60 % [Ra19], depending on the research field considered and methodology approach. However, the reasons why reproducibility fails are reasonably similar. In many cases, it is impossible to access the data and code of other researchers, which significantly decreases the chance of reproducing scientific results [La22, St19a, SSM18, CL15]. If research data is accessible, missing directions on using the given data and code is another substantial issue [Ma20, Ra19, St19a]. Therefore, we consider aiming attention to reasons why reproducibility fails as more beneficial than addressing reproducibility rates alone.

With digital research data playing an essential role in most research fields and technological advances that facilitate access to data sharing platforms, scientific journals and other stakeholders across many scientific disciplines introduced policies on data handling to improve research data sharing. The investigations by [La22, SKL18, SSM18, CL15] show that the reproducibility of scientific results correlates with the availability of research data and that the introduction of policies enforcing obligatory data sharing by scientific journals significantly affects the availability of research data. Further recommendations by the research community to increase the reproducibility of scientific results include

- mandatory code sharing and reproducibility checks by authors or publishers [La22; SKL18; CL15],
- the use of code control systems or individual training to improve code and data readability [Ra19, CP16],
- permanent contact addresses to reach out to researchers familiar with a specific publication [CP16], and
- a change in research culture that rewards code and data quality over citations and publications [St19a].

Our reproducibility study follows the approach by [St19a]. The authors use a questionnaire [St19b] to assess the reproducibility of scientific findings based on data accessibility, data availability, and the correctness of the obtained results. The study allows for systematic identification of reasons for failed reproducibility that agrees with the ACM's definition of reproducible science [ACM20]. Although we extended the questionnaire of the original study for further prospective analysis (see associated dataset), we use the same criteria to investigate the fraction of reproducible scientific publications, assess reasons for failed reproducibility, and compare our findings to the original study.

Following the criteria presented by [St19a], we investigate whether a publication includes a statement on data availability. We do not conduct a reproducibility analysis if the publication does not include a statement where the associated data can be accessed. In order to highlight publication practices in mathematics and related fields, we individually classify theoretical publications describing a new algorithm when they do not include digital artifacts and a data availability statement. Suppose a publication includes a data availability statement. In that case, we attempt to receive data from a publicly available source (e.g., an academic repository) or by contacting the authors or a third party. If the contacted persons did not reply within two weeks after sending a reminder, we did not attempt to investigate the reproducibility further. Once digital artifacts are available, we investigate whether we can reproduce the scientific findings using the given data. If the results only partly reproduce the findings, we consider the results *some reproducible* (see questionnaire in the associated dataset).

# 3 Results of an analysis over time

## 3.1 Reproducibility of results

The results of the study are summarized in Fig. 1. The figure shows why the reproducibility of the investigated publications fails according to the criteria *Data accessibility*, *Data availability*, and *Reproducibility*.

Of 108 examined papers, 26 publications are purely mathematical and investigate algorithms for data assimilation without addressing their implementation. Therefore, such publications do not involve the use of digital artifacts. Only half of the 82 publications relying on digital data include a statement on data availability. Accordingly, we merely attempted to receive the research data from 41 out of 108 publications (38 %).

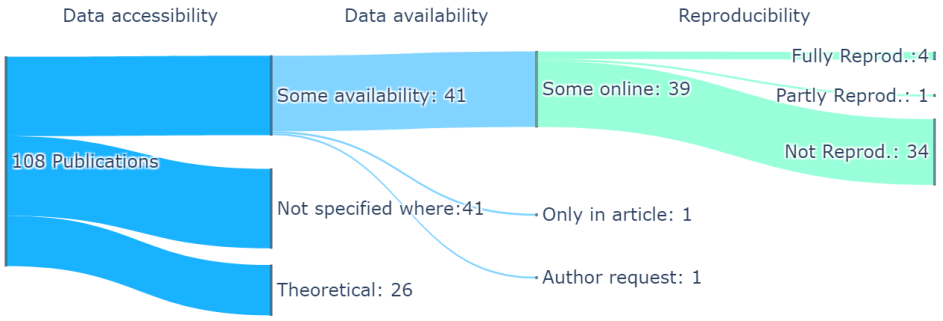


Fig. 1: Sankey diagram showing the results of the reproducibility study based on the criteria Data accessibility, Data availability, and Reproducibility. The down bends indicate at which point we discontinued the study.

We obtained digital artifacts for 39 out of 41 publications (95 %) that include a data availability statement. The data was provided via institutional homepages or data repositories, such as Zenodo, OSF, and Github, or by contacting the authors personally.

Of the 39 publications that have available artifacts, one experiment produces results that correspond to some of the findings presented in the paper but yields an index error during the generation of a plot. Since this error can likely be fixed with some additional code description, we consider this analysis partly reproducible. Overall, five out of 39 publications with data access were fully or partly reproducible (13 %). When considering all 108 publications, 4.6 % of the investigated experiments are at least partly reproducible.

The rate of reproducible research in the CRC is slightly higher than in the original study by [St19a] (see Fig. 2). Analyzing the reproducibility of their investigated publications [St19b] based on *Data accessibility*, *Data availability*, and *Reproducibility* yields that six out of 380 experiments (1.6 %) are at least partly reproducible. When comparing the reasons for discontinuing the investigation in the three respective reproducibility criteria, we see that publications in the CRC have a larger share of theoretical papers and missing data availability statements (62 % vs. 28 %). In contrast, the rate of available artifacts of publications including such statements is much higher in the CRC publications (95 % vs. 60 %). However, the scale of irreproducible scientific results is similarly prevalent despite available artifacts.

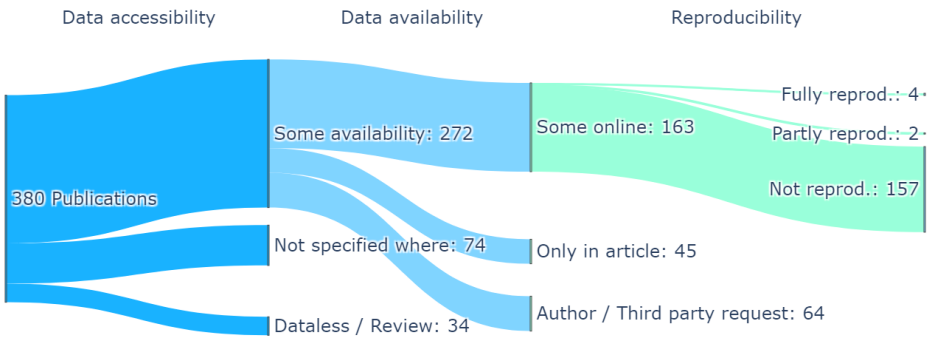


Fig. 2: Sankey diagram showing the results of the reproducibility study by [St19a]

### 3.2 Reasons for failed reproducibility

| Reason for failed reproducibility | Number of affected publications |
|---|---|
| **1. Data Accessibility** | **67 (62 %)** |
| Missing data availability statement | 41 (38 %) |
| Theoretical publication | 26 (24 %) |
| **2. Data Availability** | **2 (2 %)** |
| **3. Reproducibility** | **34 (31 %)** |
| Incomplete data | 20 (19 %) |
| Software error | 8 (7 %) |
| Insufficient documentation | 5 (5 %) |
| Proprietary software | 1 (1 %) |

Tab. 1: Reasons for failed reproducibility of 108 papers investigated in this study

Our study results enable the identification of barriers that limit the reproducibility of scientific results. While the original study [St19a] was carried out in hydrology, the CRC publications have a stronger focus on applied mathematics and related application areas. Therefore, the portion of dataless publications that describe advanced numerical approaches is naturally more prominent, leading to a relatively high number of publications labeled as non-reproducible according to our definition. Since translating such numerical techniques into computer code requires extensive mathematical knowledge [Ra19], it is inconvenient for researchers, particularly from other fields, to implement and verify the scientific findings based on the given equations in the respective publications. This circumstance alludes to a data management culture in mathematics and related fields where researchers are often not encouraged to provide digital artifacts for their algorithms.

In order to promote the sharing of research data, several scientific journals have research data policies in place that require authors to include a data availability statement. With the CRC publications, however, a substantial fraction of journals did not expect such a statement at the time of publication. At the same time, many of our researchers did not provide information on data availability on their own accord, leading to a high rate of inaccessible research data.

In contrast to the analysis by [St19a], we were able to receive at least some digital artifacts for most publications that have accessible data. We did not receive the associated data via author request in one case. In another case, the given example code in the publication was insufficient to perform the analysis. However, when artifacts were available, we observed that only in a few cases is it possible to reproduce the published research results. The reasons for failed reproducibility are rather heterogeneous but can, in principle, be classified as *missing data*, *insufficient instructions*, *software errors*, and *inaccessible software*. Most reproducibility attempts in our investigation fail due to missing data. Such failed attempts include incidents where authors did not provide all artifacts necessary to conduct the analysis or where specific resources were not available anymore, for example, because of broken links or lacking details of datasets taken from

data catalogs. Another major limitation was caused by incomplete instructions on how the associated artifacts must be used to reproduce the results. We also encountered several cases where software errors prevented further analysis of the provided artifacts. Such errors refer to weaknesses in the provided computer code, unclear dependencies, or inconsistencies with associated software libraries. In one case, the provided artifacts relied on a proprietary software library that we could not include in our investigation.

Table 1 summarizes the reasons for failed reproducibility based on the criteria *Data accessibility*, *Data availability*, and *Reproducibility*. The *Reproducibility* category includes publications where digital artifacts are at least partly available. Although there are several cases where multiple reasons account for failed reproducibility despite available artifacts (e.g., insufficient documentation and software error), we only characterize the most significant reason that caused our reproducibility investigation to be unsuccessful for each publication. Consequently, we consider the insufficient provision and documentation of digital artifacts and software errors as significant limitations in reproducing scientific results from available data. Furthermore, we attribute deficiencies in data management culture and little motivation to share data as significant limitations to accessing research data.

### 3.3 Discussion of the results

We set out to investigate and monitor the development of research data management within an interdisciplinary CRC. Our investigation demonstrates that shortcomings in data accessibility and difficulties in reproducing results from available artifacts considerably limit the reproducibility of scientific findings. This issue affects multiple research fields associated with the CRC, and several researchers made equivalent observations [La22, St19a, SKL18, SSM18, CP16]. However, recent recommendations to implement software and reproducibility tests in research [La22, SKL18, CP16] and attempts to push for open data culture in research [St19a, CL15] coincide with the community's demands to promote data accessibility and reproducibility.

Specifically, previous investigations have shown that introducing a mandatory data sharing policy by scientific journals increases research data availability and, therefore, the reproducibility of scientific results [La22, SSM18]. During the period covered by this study, various scientific journals introduced research data policies that make the sharing of associated digital artifacts mandatory. At the same time, there has been a general trend toward open science in many research communities. Accordingly, funding agencies and research institutions implemented corresponding guidelines and recommendations [De22, UP19, ERC17]. Furthermore, since the beginning of its first funding period in 2017, the CRC has provided extended personal support in research data management to scientists. Since 2017, various workshops have been conducted, and support structures have been established to assist and sensitize researchers throughout the research data lifecycle to generate reproducible research results.

We attribute these developments to the positive changes in the CRC's publications that occurred over time, as presented in Figure 3. The figure illustrates the relative annual data accessibility and reproducibility rates for the investigated publications. The data accessibility fluctuates, and we see a moderate decrease in papers that do not include a data availability statement. At the same time, the reproducibility of scientific results increased and four out of 14 investigated papers published in 2021 (29 %) are fully reproducible. Due to a significant increase in reproducible publications with a similar rate of publications sharing data over time, we attribute this progress to several community developments, such as journal policies, but also to the CRC's support structures in research data management established during this period. Therefore, we conclude that implementing advanced data management in research increase the reproducibility of scientific results.

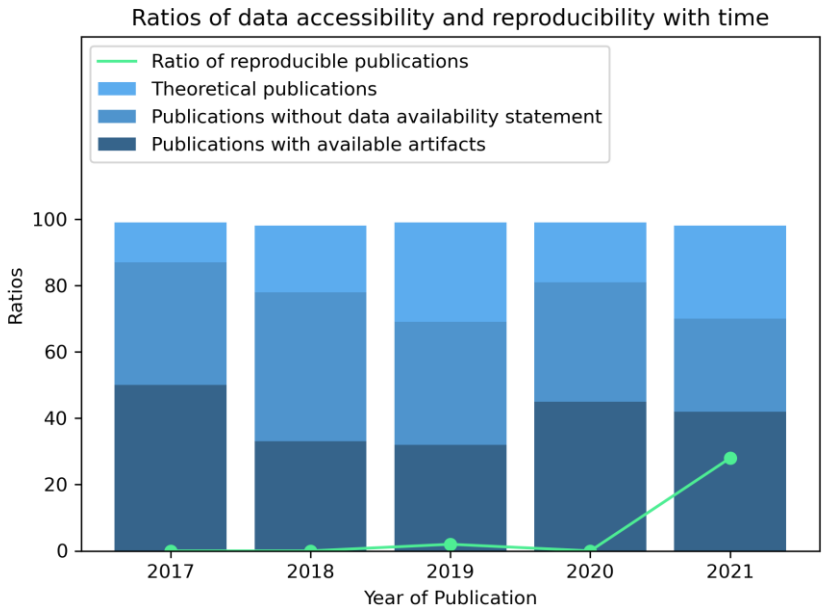


Fig. 3: Annual ratios of the investigated publication’s data accessibility and reproducibility

## 4 Summary and outlook

Our investigation indicates that significant portions of published scientific findings fail reproducibility checks due to shortcomings in data accessibility and reproducibility of results with available artifacts. Although we have positive experiences with implementing an IT infrastructure to facilitate data sharing and archiving and providing individual assistance in data management to researchers, we see further demand for support structures that promote the accessibility and reproducibility of scientific findings. These structures include support on an individual level, where researchers receive assistance in continuous reproducibility checks and training in data management to ensure adequate data availability and documentation independently.

However, it also addresses research culture in general. For example, we strongly appeal to researchers in mathematics and related fields to support their theoretical publications with computer code. Such data publications facilitate the implementation of numerical techniques and, therefore, increase the impact of numerical methods on other research fields. An interdisciplinary and convenient strategy for this purpose is the introduction of policies that make the sharing of research data mandatory since data availability correlates with reproducibility. Such policies allude to a research culture that recognizes data publications as scientific achievements. We recommend establishing structures that value data quality alongside publications and citations when assessing research quality.

In summary, in order to improve the reproducibility of research results, we recommend to

- establish institutional and nationwide structures that support scientists in data sharing and research data management,
- apply compulsory artifact sharing and reproducibility checks by authors, publishers, and research institutions,
- use version control systems and provide training to researchers to improve the comprehensibility of data and code,
- recognize and reward research data as scientific achievements, including academic hiring processes.

## Acknowledgements

This work was funded by Deutsche Forschungsgemeinschaft (DFG) - Project-ID 318763901 - SFB1294. The data supporting the findings of this study are available at https://doi.org/10.5281/zenodo.6543497.

## Bibliography


[ACM20] Association for Computing Machinery: Artifact Review and Badging Version 1.1, https://www.acm.org/publications/policies/artifact-review-and-badging-current, accessed: 13/05.2022.

[Ba16] Baker, M.: 1,500 scientists lift the lid on reproducibility. Nature 533, 452–454, 2016. https://doi.org/10.1038/533452a

[CL15] Chang, A. C.; Li, P.: Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say ”Usually Not”. Finance and Economics Discussion Series 2015-083. Board of Governors of the Federal Reserve System, Washington, 2015. http://dx.doi.org/10.17016/FEDS.2015.083

[CP16] Collberg, C.; Proebsting, T.A.: Repeatability in computer systems research. Commun. ACM, 59/3, 62–69, 2016. https://doi.org/10.1145/2812803

[De22] Deutsche Forschungsgemeinschaft: Guidelines for Safeguarding Good Research Practice. Code of Conduct, Version 2, 2022. https://doi.org/10.5281/zenodo.6472827

[ERC17] European Research Council: Guidelines on Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020, https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf, accessed: 13/05/2022.

[EU19] EU Directive 2019/1024 on open data and the re-use of public sector information, 2019. http://data.europa.eu/eli/dir/2019/1024/oj, accessed: 13/05/2022.

[Ge21] Geßner, H.: Transparently Safeguarding Good Research Data Management with the Lean Process Assessment Model. E-Science-Tage - Share Your Research Data, 2021. https://doi.org/10.11588/heidok.00029719

[Gö19] Götzelmann, G.; Hegel, P.; Krewet, M.; Söring, S.; Tonne, D.: Aspects of Digital Infrastructures - Challenges and Perspectives of Research Data in the Collaborative Research Centre “Episteme in Motion”. BIBLIOTHEK – Forschung und Praxis, 2019, AR 3284.

[KS13] Kindling, M; Schirmbacher, P.: „Die digitale Forschungswelt“ als Gegenstand der Forschung: Lehrstuhl Informationsmanagement. Wissenschaft & Praxis 64(2-3), pp. 127-136, 2013. https://doi.org/10.1515/iwp-2013-0017

[La22] Laurinavichyute, A; Yadav, H.; Vasishth, S.: Share the code, not just the data: A case study of the reproducibility of articles published in the Journal of Memory and Language under the open data policy. PsyArXiv Preprint, 18 Apr 2022. https://doi.org/10.31234/osf.io/hf297

[Ma20] Maassen E.; van Assen, M.A.L.M.; Nuijten M. B; Olsson-Collentine A.; Wicherts J. M.: Reproducibility of individual effect sizes in meta-analyses in psychology. PLOS ONE 15(5), e0233107, 2020. https://doi.org/10.1371/journal.pone.0233107

[Nü18] Nüst, D.; Granell, C.; Hofer, B.; Konkol, M.; Ostermann, FO.; Sileryte, R.; Cerutti, V.: Reproducible research and GIScience: an evaluation using AGILE conference papers. PeerJ 6:e5072, 2018. https://doi.org/10.7717/peerj.5072

[Ob20] Obels, P.; Lakens, D.; Coles, NA.; Gottfried, J.; Green, SA.: Analysis of open data and computational reproducibility in registered reports in psychology. Adv. Methods Pract. Psychol. Sci. 3, 229–237, 2020. https://doi.org/10.1177/2515245920918872

[Pe17] Perrier, L.; Blondal, E.; Ayala, AP.; Dearborn, D.; Kenny, T.; Lightfoot, D. et al.: Research data management in academic institutions: A scoping review. PLoS ONE 12(5): e0178261, 2017. https://doi.org/10.1371/journal.pone.0178261

[Ra19] Raff, E.: A step toward quantifying independently reproducible machine learning research. Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, pp. 5485-5495, 2019. https://doi.org/10.48550/arXiv.1909.06674

[SKL18] Stockemer, D.; Koehler, S.; Lentz, T.: Data Access, Transparency, and Replication: New Insights from the Political Behavior Literature. PS: Political Science & Politics 51(4), 799-803, 2018. https://doi.org/10.1017/S1049096518000926

[SSM18] Stodden, V.; Seiler, J.; Ma, Z.: An empirical analysis of journal policy effectiveness for computational reproducibility. Proc. Natl. Acad. Sci. 115, 2584–2589, 2018. https://doi.org/10.1073/pnas.1708290115

[St19a] Stagge, J.; Rosenberg, D.; Abdallah, A. et al.: Assessing data availability and research reproducibility in hydrology and water resources. Sci Data 6, 190030, 2019. https://doi.org/10.1038/sdata.2019.30

[St19b] Stagge, J.; Abdallah, A.M.; Rosenberg, D.E.: jstagge/reproduc_hyd: Repository as published in Scientific Data (2.0). Zenodo, 2019. https://doi.org/10.5281/zenodo.2562268

[UP19] Universität Potsdam: Forschungsdaten-Policy der Universität Potsdam, Auszug aus den Amtlichen Bekanntmachungen Nr. 18 vom 30.9.2019, https://www.uni-potsdam.de/fileadmin/projects/ambek/Amtliche_Bekanntmachungen/2019/ambek-2019-18-1361-1362.pdf, accessed 13/05/2022.

[WGS16] Wang, WM.; Göpfert, T.; Stark, R.: Data Management in Collaborative Interdisciplinary Research Projects - Conclusions from the Digitalization of Research in Sustainable Manufacturing. ISPRS Int. J. Geo-Inf. 2016, 5, 41. https://doi.org/10.3390/ijgi5040041

[WC19] Wiljes, C.; Cimiano, P.: Teaching Research Data Management for Students. *Data Science Journal*, 18(1), 38, 2019. http://doi.org/10.5334/dsj-2019-038