Emerging Roles and Responsibilities of Libraries in Support of Reproducible Research

Birgit Schmidt

State and University Library, University of Göttingen, Göttingen, Germany, bschmidt@sub.uni-goettingen.de, orcid.org/0000-0001-8036-5859

Andrea Chiarelli

Research Consulting, Nottingham, UK, andrea.chiarelli@research-consulting.com, orcid.org/0000-0001-7336-8330

Lucia Loffreda

Research Consulting, Nottingham, UK, lucia.loffreda@research-consulting.com, orcid.org/0000-0002-3548-3124

Jeroen Sondervan

Utrecht University, Utrecht; Dutch Research Council (NWO), The Hague, The Netherlands, j.sondervan@nwo.nl, orcid.org/0000-0002-9866-0239

Abstract

Ensuring the reproducibility of research is a multi-stakeholder effort that comes with challenges and opportunities for individual researchers and research communities, librarians, publishers, funders and service providers. These emerge at various steps of the research process, and, in particular, at the publication stage. Previous work by Knowledge Exchange highlighted that, while there is growing awareness among researchers, reproducible publication practices have been slow to change. Importantly, research reproducibility has not yet reached institutional agendas: this work seeks to

highlight the rationale for libraries to initiate and/or step up their engagement with this topic, which we argue is well aligned with their core values and strategic priorities. We draw on secondary analysis of data gathered by Knowledge Exchange, focusing on the literature identified as well as interviews held with librarians. We extend this through further investigation of the literature and by integrating the findings of discussions held at the 2022 LIBER conference, to provide an updated picture of how libraries engage with research reproducibility. Libraries have a significant role in promoting responsible research practices, including transparency and reproducibility, by leveraging their connections to academic communities and collaborating with stakeholders like research funders and publishers. Our recommendations for libraries include: i) partnering with researchers to promote a research culture that values transparency and reproducibility, ii) enhancing existing research infrastructure and support; and iii) investing in raising awareness and developing skills and capacities related to these principles.

Keywords: research reproducibility; open science; publishing; scholarly communication; research data management; FAIR; code and software; research support; research libraries.

1. Introduction

On the one hand, research libraries are a stable entity when it comes to reliable information management for their institution and broader research communities. On the other hand, they are undergoing a constant transformation based on their changing environments as well as related roles and responsibilities for example in the area of open research. Based on continually emerging opportunities and challenges, libraries adopt and drive new or enhanced information practices. In recent years, the extent to which research findings can be trusted has been challenged based on the inability to reproduce or replicate research findings (the so-called "reproducibility crisis"),[1] which is often due to limited or no access to research data, lack of documentation in methodologies and/or limited or no availability of the software and code used to run an analysis. These issues, which are often considered as part of research integrity (Diaba-Nuhoho & Amponsah-Offeh, 2021), apply to published research, and can severely affect a researcher's career, ranging from minor errors to severe ones leading to retractions or even investigations of research misconduct.

We acknowledge that there is no universally agreed definition of "research reproducibility", that is the related terms "replication", "reproduction" and

"repetition" are sometimes used interchangeably or as distinct concepts, depending on context and research discipline (Stanford Encyclopedia of Philosophy, 2018; National Academies of Sciences, Engineering, and Medicine [NASEM], 2018, 2019). For the purpose of this paper, we will refer to reproducibility as the ability to reproduce the findings of a study based on access to the original data and methods/procedures. This usually requires a detailed description of the methods used to process and analyse the data, access to any relevant datasets (and related documentation) and an ability to obtain and run computer code, where appropriate. It must be noted that definitions of reproducibility often focus on "computational reproducibility". Findings based on qualitative methodologies (e.g. literature reviews, interviews, manual thematic analysis) will often not be fully reproducible based on subjective elements in assessment, classification, selection, contextualisation or interpretation processes.

Between 2020 and 2021, Knowledge Exchange (KE) investigated the publication of reproducible research outputs, with support from a group of international open science experts. The KE partners are six key national organisations within Europe tasked with developing infrastructure and services to enable the use of digital technologies to improve higher education and research: IT Center for Science (CSC) in Finland, Centre national de la recherche scientifique (CNRS) in France, the Danish e-Infrastructure Consortium (DeiC) in Denmark, the German Research Foundation (DFG) in Germany, Jisc in the UK and SURF in The Netherlands. The study was initiated by KE's Open Access Expert Group, which focuses on the ongoing transformation of the scholarly publishing landscape.[2] One core finding was that ensuring the reproducibility of research is a multi-stakeholder effort, with challenges and opportunities for individual researchers and research communities, librarians, publishers, funders and service providers. These emerge at various steps of the research cycle, and, in particular, at the publication stage.

The output of this work was a public report (Chiarelli et al., 2021a, 2021b), which was based on an extensive literature review of almost 130 sources and a mix of interviews and focus groups with 51 stakeholders from 12 countries. In this landscape study, stakeholders were considered at the micro, meso and macro level based on the *KE Open Scholarship Framework* (Neylon et al., 2019, cf. Figure 1).
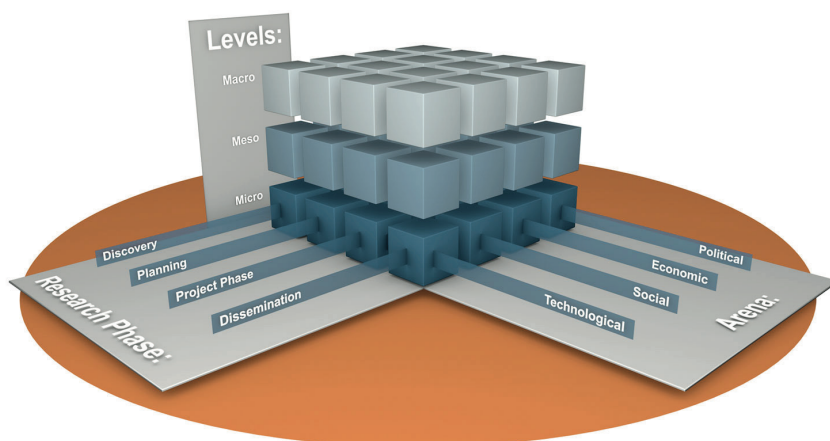
Across its levels, the Open Scholarship Framework considers researchers and research groups, their disciplinary communities and institutions as the content creators (micro level); academic libraries and their institutions as well

as publishers and infrastructure providers in supporting and strategic roles (meso level); and research funders and policymakers as those who influence the broader ecosystem and establish recommendations and requirements (macro level).

In this context, the exact role that librarians should or could play remains unclear. For example, why should librarians care about the reproducibility of (published) research findings? Through what mechanisms could they enable and foster best practices in research and publishing, leveraging the services and activities that they manage in collaboration with other stakeholders? In this paper, we have sought to translate the results of the KE study into practical tools for research libraries.

Our paper begins with a background section that provides context and foundational information on research reproducibility, including with regard to the role of libraries. This is followed by an analysis of stakeholder roles and key touchpoints for librarians, enriched with relevant quotes. The paper ends with a summary of our findings and a series of recommendations to enhance the role of research libraries in the reproducibility landscape.

Fig. 1: Knowledge Exchange Open Scholarship Framework.

2. Background

Libraries' engagement with reproducibility as an activity and set of skills has multiple perspectives: 1) From the perspective of *collections management*, libraries select, manage, curate and ensure the discoverability of the published record of research outputs. In the academic and societal environment in which libraries operate, challenges such as the "reproducibility or replication crisis" can affect the trust in these collections. 2) In the context of *research support*, libraries directly engage with researchers by providing support and infrastructures for research data management as well as training in data (science) literacy and its relationship to open science (e.g. FAIR and open data) (Lyon, 2016; Quan, 2021). Building up services for the management and curation of research software is a relatively new yet emerging responsibility for libraries (Chassanoff et al., 2018; Martinez-Ortiz et al., 2023). 3) Institutional *publishing services* develop publication policies, workflows and infrastructures and provide support for transparency and openness (e.g. Nosek et al., 2023). This includes guidance on data publishing and citation, e.g. via own or external repositories, availability statements and licensing. 4) Across all these service areas, libraries contribute to awareness raising, training and implementation support for *responsible research practices* at various stages of the research and scholarly communication cycle as a safeguard for reliable research.

Over the last few years, the research reproducibility landscape has been rapidly evolving, as discussions around open access, FAIR and open data, research integrity, open source and open infrastructures continue to converge. More specifically, it has been pointed out that open science provides a means to improve the quality, transparency and reproducibility of research (Munafò et al., 2017) leading to more accountability of research. This also holds for the step of publishing research outcomes: "executable research papers" support reproducibility through a combination of text, raw data and code that (re)creates the analysis and figures, and offers readers means to interact with or manipulate the code (Tsang & Maciocci, 2020). For researchers, taking up this approach for writing papers can turn out as an insightful learning experience and journey through open science (Lasser, 2020). The approach is however far from becoming the established norm as this would require a cultural shift of the scientific community towards making data and code openly available and a recognition of the skills and efforts needed to implement all necessary steps.

To support and foster reproducibility across the research cycle, many stakeholders including researchers, scholarly societies, funders and publishers have created guidelines and recommendations (e.g. KNAW, 2018; Munafò et al., 2017; National Institutes of Health, 2014; Nosek et al., 2023; Wellcome Trust, 2017).[3] Moreover, collaborative initiatives led by researchers have formulated documentation standards in support of reproduction (e.g. Project TIER Protocol),[4] hands-on guides (e.g. Turing Way Handbook[5] (The Turing Way Community, 2022)), reviews of existing infrastructures (e.g. Konkol et al., 2020), assessed the reproducibility of research results in funding programmes (e.g. European Commission et al., 2022a), and are co-creating activities and policy recommendations (e.g. Ross-Hellauer et al., 2022). These documents provide guidance for researchers but also point out infrastructures and tools as well as support services that are provided by a range of stakeholders. From a conceptual perspective Lyon (2016) proposed a 3-dimensional model or spectrum for open science, which includes the 3rd axis of transparency. In this model librarians act as 'transparency agents' that advocate, promote and demonstrate particular behaviours and good practices throughout the research lifecycle, which over time will lead to cultural change towards a more open science environment. Sayre and Riegelman (2019) compared a selected set of reproducibility guidelines and have mapped libraries' expertise and services to key themes of these guidelines. LaPolla et al. (2022) investigated motivations, approaches and lessons learned from libraries' engagement – typically in collaboration with other departments – in teaching rigor and reproducibility in the health sciences.

When it comes to reproducibility, research libraries, publishers and other stakeholders need to navigate a web of interconnected roles and responsibilities, which includes not only the advice and support they traditionally provide, but also an ever-broadening range of tools and workflows that span from choosing a repository for the deposit of data and code to preparing research compendia[6] or executable research articles (Packer, 2020). Positively, research libraries have taken up innovative new roles, such as the "data librarians/ stewards" that are becoming more common across Europe – in particular in the context of implementing the FAIR data principles (Wilkinson et al., 2016) – and the "reproducibility librarians" that are emerging in the USA and operate at the interface between the various facets of open science (Steeves, 2017).

The role of libraries and research support services with regard to reproducibility is also changing in terms of digital infrastructures. Researchers require

more and more support in navigating the fast-paced world of technological solutions emerging to support open science practices, thus leading to uncertainty and a constant need to upskill both the research base and support staff. Whether internally or in partnership with external providers, research libraries also continue to develop institutional and thematic repositories, which can be interconnected with a web of digital objects to enable reproducible research to take place. The European Open Science Cloud (EOSC) is an example of ongoing efforts to build a trusted distributed digital research infrastructure that supports verifiable and reproducible research (European Commission, 2022b).

Finally, the broader context of the COVID-19 pandemic has highlighted the role of research reproducibility in enabling trust in the published record. For example, it was found that rapidly emerging preprints tackling important public health questions didn't always meet high transparency criteria and could improve their data sharing and scientific reporting practices (Sumner et al., 2020). Research libraries can play a dual role in this regard, as they can be both the custodians of research objects and advisers to local academics who seek to work in a reproducible way and enhance the quality of their institution's publications.

3. Methods and Data

This work presents a secondary analysis of the findings that were gathered during the 2020–2021 KE exercise on the Art of Publishing Reproducible Research Outputs.[7] This study included a literature review, stakeholder consultation and was presented at the 2022 LIBER conference.

To provide an updated view of the reproducibility landscape, we have built on the literature review conducted as part of the KE exercise as a starting point. This literature review included almost 130 literature sources from across academic disciplines and multidisciplinary sources and provided an overview of definitions, problems, recommendations and solutions or initiatives in the context of reproducibility (Chiarelli et al., 2021a; Loffreda & Chiarelli, 2021). In the present study, we have revisited the literature findings to focus particularly on the key roles and responsibilities for research libraries. We have then conducted a brief review of literature to reflect any recent developments in the field of reproducibility that may

affect or relate to these roles, such as the establishment of new initiatives and networks.

In addition to the updated review of literature, a secondary analysis of the interview transcripts gathered as part of the original exercise (Chiarelli et al., 2021b) has been conducted. Transcripts that have been reviewed include interviews with 41 international stakeholders, including research librarians, researchers, publishers and funders, and focus groups with a further 10 infrastructure providers. Text searches were conducted on these sources, including searches for "library" and "steward", to identify examples of those in research support roles supporting the publication of reproducible research. Quotes from interviews are used throughout this paper to illustrate our key findings.

Finally, based on the above-mentioned methodology, we have considered how the findings of the updated landscape analysis align to LIBER's 2023–2027 strategy (Frey et al., 2022). This approach has informed the development of the three main recommendations presented in the final section of this paper.

4. Limitations

We note that our updated literature and landscape analysis has been conducted during January and February 2023, thus representing a snapshot of the reproducibility landscape at a given point in time. It should also be noted that the original analysis has been underpinned by thematic coding. By nature, this relies on the subjective interpretation of the researcher.

5. Stakeholder Roles and Key Touchpoints for Librarians

Due to the multifaceted nature of research reproducibility, roles and responsibilities are inevitably complex and must be shared between different stakeholder groups at the aforementioned micro (researchers and disciplines), meso (institutions, libraries, publishers, software providers) and macro levels (research funders, policymakers). In particular, we note that cooperation is essential throughout the research lifecycle, as the publication of a reproducible research output requires a range of considerations from the inception of

a research project all the way through its delivery and dissemination. One academic interviewee highlighted this point:

> *"I think the main thing is that these things need to be thought of early in the process, and if you start thinking about them too late it becomes harder to do them. So, in the context of data sharing for example, if you're collecting data from human participants you need to ask them for consent for data sharing, and if you leave it until after you've collected the data you can't share the data as openly as if you have collected consent from people early on."* (Academic interviewee 1)

One of the key challenges in research reproducibility is the need to share a broad range of research objects (as applicable) alongside one's outputs, including detailed methods, pre-registrations, research data, any relevant code/software and information about the software environment used. Notably, in listing these research objects we do not mean that reproducibility is inherently quantitative or computational in nature. It is, however, essential to acknowledge that the terminology we employed above may be misunderstood or misinterpreted in some disciplines, which highlights the importance of stakeholders speaking a shared language. For example, reproducibility guidance may be inappropriate at the funder or publisher level and much more helpful at the funding council or journal level, so as to add an extent of tailoring that will allow the target audience to understand any requirements. This sentiment was also expressed by another academic interviewed as part of the landscape study:

> *"I think a recognition of the fact that there's no single bullet, there will not be one system, or one set of practice that is going to rule everything, in that everybody will have to comply with. So that the recognition that it has to be, there will be different processes, different tools, and so it's essentially a training challenge, and innovation should come in the area of education."* (Academic interviewee 2)

When it comes to the sharing of the above-mentioned research objects, it is clear that researchers play a key role: they are responsible for designing, delivering and disseminating the research and hold the information on relevant research objects (including metadata). At the same time, researchers can act as peer reviewers in cases where journals have reproducibility requirements, meaning that they may be responsible for checking that someone else's work is reproducible. In the case of reproducibility, it is also helpful to think of researchers as members of disciplinary communities: in most cases,

reproducibility is unlikely to become a widespread concern unless the (sub) discipline establishes that this should be the case. For example, some disciplines show widespread data and code sharing (with implications on reproducibility), particularly where physical or digital infrastructure is shared (e.g. particle physics, astronomy), while others are currently not engaging with the concept. It is, however, recognised that reproducibility "is relevant across scientific disciplines and cultures, be it in the humanities, engineering, life science, natural sciences, or social sciences", although projects in different disciplines may take "demonstrably different approaches, varying from enhancing step-by-step reproducibility through code-based transparency to tracing the origin of an argument through publication lines" (Rahal et. al., 2022). Furthermore, the role of individual reproducibility champions (including with the support of national reproducibility networks) in raising the profile of practices underpinning reproducible publication in disciplines where these remain uncommon is broadly recognised (Chiarelli et al., 2021b; Ross-Hellauer et al., 2022). This is where the potential for institutions to make a difference is highest, as fostering such efforts and providing related support is conducive to the broader uptake of new practices. As one academic interviewee highlighted:

> *"I think right now you have a few champions per community, which are driving efforts at different pace, in different spaces. So you have the people who are involved in producing tools… there are people who have driven studies reproducing large numbers of results. So I think it's more a bunch of individuals and what we are trained to do is to find a way to get a forum to get some of those to communicate and inspire others."* (Academic interviewee 3)

Arguably, libraries are uniquely placed to provide support around research reproducibility, as they are the institutional reference point for both authors and publishers and are typically well connected with other professional services such as IT, software engineers, ethics, data governance and data protection. Clearly, the remit of librarians varies based on national and institutional contexts, meaning that the provision around research reproducibility can take a range of different forms. For example, New York University and Oregon Health & Science University provide dedicated research data management and reproducibility librarians,[8] while at the Delft University of Technology reproducibility is within the remit of research data stewards (Martinez-Lavanchy et al., 2022). In other cases, reproducibility may be supported by a mix of roles, including research object curators or subject librarians, as was suggested by a research support librarian:

> *"I mean it might be a bit difficult because there are so many different disciplines and fields. I was thinking whether that would be somehow related to the library services, that's the natural place, where there could be someone who they could ask to help. Like if you want to have some kind of geometric study you would just go to your library staff and they could help you. It could be something similar with reproducible science."* (Librarian interviewee 1)

A wide range of initiatives is being introduced as part of grassroots efforts to support research reproducibility, and these are often advertised and signposted within institutions by libraries in collaboration with researchers, IT and research services. These include community-initiatives such as ReproHacks,[9] ReproducibiliTea,[10] The Carpentries[11] and the Turing Way Handbook[5] (The Turing Way Community, 2022) and the national Reproducibility Networks (e.g. UK, Germany, Sweden, Portugal)[12] as well as local awareness raising and capacity building efforts (e.g. Rethlefsen et al., 2018). These initiatives are complemented by platforms, software and technical tools to support the management, sharing and execution of relevant research objects, and a list of examples is available in Chiarelli et al. (2021b). At the high level, we note the importance of tools for reproducibility checking, software containerisation, code sharing and documentation, reproducible environments, research compendia and more.[13] In most cases, individual researchers will not be familiar with these tools: given the great and continually increasing pressures on researchers' time, it is essential for institutions (and, in most cases, libraries) to provide an extent of training and/or support around these systems if reproducibility is considered as a strategic priority. One institutional librarian noted:

> *"We do our best to get that information out, but there are still a lot of things to do I believe, to try to teach these skills also. Not only the tools and support, but also for them to be able to use these tools."* (Librarian interviewee 2)

In practice, research libraries are only beginning to engage with reproducibility, including because this is not currently a priority in all disciplines and is lagging behind other portions of the open research agenda (e.g. open access, open data), partly due to the mix of complex practical requirements and expectations. This is where the role of research funders and government players becomes apparent. To date, minimal pressure has been perceived around reproducibility at the policy level, and institutions have consequently not seen this to be an area for strategic investment in most cases. This is not due to a lack of interest in reproducibility, but rather in

acknowledgement that library (and, more broadly, institutional) budgets continue to be under growing pressure and areas of expenditure have to be prioritised. An interview with an institutional librarian touched on this topic, and they noted:

> *"That's [reproducibility] something we are actively trying to support but we get handcuffed in our own jobs because of lack of resources and lack of funding. People would rather put money into new science than preserving old science and that's also been a huge problem. So we talk a big game about reproducibility but no one's thinking in the long term."* (Librarian interviewee 2)

As noted above, publishers in some disciplines are beginning to introduce reproducibility requirements, however, meaning that strategic pressures that might often arise at the macro level are actually emerging at the meso level and as part of micro-level grassroots movements led by researchers and disciplines

6. Conclusions

This paper set out to identify the rationale for libraries to engage with the reproducibility of (published) research findings and to outline a range of mechanisms and practices that would enable this. The discussion so far indicates that libraries have a natural role to play in raising awareness about responsible research practices – including efforts that aim at increasing transparency and enhancing reproducibility of research – and for providing support for their day-to-day implementation. For this, libraries can build on their good connections to academic communities and are well aware of disciplinary needs. To cover all steps of the research and scholarly communication cycle, collaboration with other stakeholders such as research funders and publishers is essential.

In sum, we would like to propose three recommendations for the the role of research libraries in the reproducibility landscape: librarians should team up with researchers (and other stakeholders) for the promotion of an (institutional) *research culture that supports and rewards* the principles of research transparency and reproducibility; build on and expand existing *research infrastructure and support*; and engage and invest in raising *awareness and* the creation of *skills and capacities*.

6.1. Institutional Research Culture

Transparency, openness and reproducibility are not just good practice principles but means to enhance the quality, integrity and effectiveness of research. To foster a research culture that encourages, acknowledges and rewards researchers' efforts in this regard, librarians should step up their engagement and act as allies and partners in the ongoing reform of research assessment (CoARA, 2022; Council of the European Union, 2022, 2023; European Commission, 2021). In particular, they can:

– Lead or contribute to the development of amendments of institutional research policies (e.g. with reference to funder policies, general or disciplinary good practice principles).
– Engage in discussion on incentives and recognition of achievements with respect to transparency and reproducibility of research (e.g. for data reuse, evaluation criteria).

6.2. Research and Publication Infrastructures, Guidance and Support

Research can be made more robust and reproducible through careful planning, management and documentation of workflows, covering all relevant inputs, processing steps, outcomes and interconnections. While there are multiple environments and tools for implementing such workflows, a crucial step is the integration with reliable research and publication infrastructures (e.g. data repositories, library-based publishing) that support the sharing, publication and archiving of research outcomes. In this context, librarians can:

– Build on existing expertise in terms of providing support, training and creating infrastructures for open access and research data management.
– Preserve and share methodology, data, software and other inputs and outputs to enable research replication and to build public knowledge (Library Publishing Coalition Ethical Framework Task Force, 2023).
– Invest in open infrastructures in the areas of FAIR data, code and software that deliver storage, tools, services and workflows for researchers to enhance their reproducibility. An essential element to take into account is the interoperability of such infrastructure.
– Engage in the management and curation of research data and research software is key to achieve greater FAIRness and long-term usability (Arguillas et al., 2022; Martinez-Ortiz et al., 2023).

- Invest in (inter)national networks that can be supportive to reproducibility (e.g. (national) Research Integrity networks, Research Data Alliance (RDA), and Research Software Alliances (ReSA)).
- Engage with open source projects and communities (Quan, 2022), as these develop the tools and practices that enable reproducible research (e.g. Project Jupyter community, ROpenSci community).[14]

6.3. Awareness, Skills and Capacities

Librarians are well-positioned to train and provide guidance on essential aspects of research reproducibility, and can integrate these into existing or emerging efforts. This touches on various steps of the research cycle, in particular:

- basic data science skills (e.g. Lyon, 2016; Oliver et al., 2019) including tidy data, data cleaning, data visualization via diverse tools (e.g. The Carpentries modules on data and software);
- research data management (e.g. Lyon, 2016; Oliver et al., 2019): data management plans, documentation and sharing of data and code, FAIR and open data;
- publishing: data and software availability statements, basics of peer review, data citation, publication formats (data or software papers, etc.) and publication venues (e.g. innovative publication platforms) that support reproducibility;
- reproducibility hackathons (repro-hacks), in collaboration with researchers (Hettne et al., 2020);
- guidance on tools that help to make research more reproducible (Lyon, 2016).

Moreover, these new roles can be made visible by adding support for research reproducibility to job descriptions, e.g. of data stewards or open science advisors and community managers. However, this also requires commitment from libraries and library communities to invest further in activities such as shared educational programmes and practical toolkits and guidance on transparency and reproducibility of research.

These aims and actions also link very well to libraries' core values and strategic priorities: libraries as engaged and trusted hubs are committed

to i) advancing research and scholarship that is increasingly equitable and open, ii) providing state-of-the-art and scholar-focused services and iii) upskilling the library workforce. The primary aim is to address the needs of their user communities, in collaboration with local, national and international partners (Association of Research Libraries [ARL], 2022; LIBER, 2022).

However, while there are good reasons for libraries to engage in activities in support of research reproducibility we would also like to point out several challenges and barriers. Research methodologies and tools have become more data-intense across almost all disciplines. As already pointed out, definitions of research reproducibility and views on its relevance vary by discipline. Librarians are on the one hand knowledgeable in the information needs of the disciplines, however, that does not necessarily include an up-to-date in-depth familiarity with the research questions and methodologies. They might therefore not be well-equipped to provide advice on all necessary steps and available options related to research reproducibility. A further challenge is to shift or secure additional resources and to address long-term sustainability of infrastructures in support of research reproducibility and open research. In particular in a fragmented landscape libraries are not necessarily recognised in their own institution as crucial players and providers of these infrastructures and services, addressing generic and disciplinary needs.

We think that research libraries are an essential part of the puzzle when it comes to achieving research reproducibility. It may vary greatly how research communities, universities and national research agenda's address the reproducibility challenge. However, we hope that with these recommendations and highlighting the libraries as a key stakeholder group, the research library community will step up their engagement for aligning their strategies and common goals in this area.

As discussed above, collaboration with other stakeholders (funders, publishers, service providers, researchers and their networks) on the local and (inter)national level is crucial to reach consensus on priorities and required investments, to provide guidance and support for implementation, to track progress made and to address emerging needs, from an overarching and where needed disciplinary perspective.

Note from the Authors

The quotes used in this article have been edited by us only to remove grammatical errors and improve readability. No adjustments were made to the content.

Data Availability

A data management plan, all datasets and code that have been created in the context of the Publishing Reproducible Research Outcomes (PRRO) project are available via Zenodo: https://zenodo.org/communities/ke-prro/. In particular, this includes datasets related to the literature review, conducted interviews and the analysis of social media data.

Acknowledgement

This article builds on the previous work and research done by Knowledge Exchange and Research Consulting. The authors would like to thank all involved colleagues and organisations for all their contributions, reflected in the final report *The Art of Publishing Reproducible Research Outputs: Supporting emerging practices through cultural and technological innovation* and which has been the essential groundwork for this endeavor.

References

Arguillas, F., Christian, T.-M., Gooch, M., Honeyman, T., & Peer, L. (2022). *10 Things for Curating Reproducible and FAIR Research* (Version 1.1). Research Data Alliance. https://doi.org/10.15497/RDA00074

Association of Research Libraries (ARL). (2022). *ARL's 2023-2026 Action Plan.* https://www.arl.org/who-we-are/#section-action-plan

Chassanoff, A., AlNoamany, Y., Thornton, K., & Borghi, J. (2018). Software curation in research libraries: Practice and promise. *Journal of Librarianship and Scholarly Communication, 6*(1), Article eP2239. https://doi.org/10.7710/2162-3309.2239

Chiarelli, A., Loffreda, L., & Johnson, R. (2021a, April 8). *Publishing reproducible research outputs - Literature findings.* [Conference presentation]. Literature findings. https://doi.org/10.5281/zenodo.4675457

Chiarelli, A., Loffreda, L., & Johnson, R. (2021b). *The Art of Publishing Reproducible Research Outputs: Supporting emerging practices through cultural and technological innovation.* Knowledge Exchange. https://doi.org/10.5281/zenodo.5521077

Coalition on Advancing Research Assessment (CoARA). (2022). *Agreement on reforming research assessment.* https://coara.eu/agreement/the-agreement-full-text/

Council of the European Union. (2022). *Research assessment and implementation of Open Science.* https://www.consilium.europa.eu/media/56958/st10126-en22.pdf

Council of the European Union. (2023). *High-quality, transparent, open, trustworthy and equitable scholarly publishing.* https://data.consilium.europa.eu/doc/document/ST-9616-2023-INIT/en/pdf

Diaba-Nuhoho, P., & Amponsah-Offeh, M. (2021). Reproducibility and research integrity: The role of scientists and institutions. *BMC Research Notes*, *14*(1). https://doi.org/10.1186/s13104-021-05875-3

European Commission. Directorate General for Research and Innovation. (2021). *Towards a reform of the research assessment system: Scoping report.* Publications Office. https://data.europa.eu/doi/10.2777/707440

European Commission. (2022a). *Assessing the reproducibility of research results in EU Framework Programmes for Research: Final report.* Publications Office. https://data.europa.eu/doi/10.2777/186782

European Commission. (2022b). *Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC).* Publications Office of the European Union. https://data.europa.eu/doi/10.2777/935288

Frey, J., Roche, J., Kautonen, H., Schmidt, B., Tsakonas, G., van Wijngaarden, H., Smit, A., Cassafieres, C. S., Burman, L., Verheusen, A., Knigge, R., Joss-Bethlehem, E., & van der Graaf, M. (2023). *LIBER Strategy 2023-2027.* LIBER. https://doi.org/10.5281/zenodo.7696568

Hettne, K., Proppert, R. K. K., Nab, L., Saunero, L. P. R., & Gawehns, D. (2020). *ReprohackNL 2019: How libraries can promote research reproducibility through community engagement.* SocArXiv Papers. https://doi.org/10.31235/osf.io/6f4zv

Konkol, M., Nüst, D., & Goulier, L. (2020). Publishing computational research—A review of infrastructures for reproducible and transparent scholarly communication. *Research Integrity and Peer Review*, *5*(10), 1–8. https://doi.org/10.1186/s41073-020-00095-y

KNAW. (2018). *Replication studies – Improving reproducibility in the empirical sciences.* Amsterdam, KNAW.

LaPolla, F. W. Z., Bakker, C. J., Exner, N., Montnech, T., Surkis, A., & Ye, H. (2022). Rigor and reproducibility instruction in academic medical libraries. *Journal of the Medical Library Association*, *110*(3), Article 3. https://doi.org/10.5195/jmla.2022.1443

Lasser, J. (2020). Creating an executable paper is a journey through Open Science. *Communications Physics, 143*(3), 1–5. https://doi.org/10.1038/s42005-020-00403-4

Library Publishing Coalition Ethical Framework Task Force. (2023). *An Ethical Framework for Library Publishing, Version 2.0*. Library Publishing Coalition https://doi.org/10.5703/1288284317619

Loffreda, L., & Chiarelli, A. (2021). *Publishing Reproducible Research Outputs—Literature sources and Thematic coding* (Version 1) [Data set]. Zenodo. https://doi.org/10.5281/ZENODO.4748748

Lyon, L. (2016). Transparency: The emerging third dimension of Open Science and Open Data. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 25(4), 153–171. https://doi.org/10.18352/lq.10113

Martinez-Lavanchy, P. M., van Schöll, P., & Zormpa, E. (2022). *TU Delft Research Data Management 101 course*. Zenodo. https://doi.org/10.5281/ZENODO.6325919

Martinez-Ortiz, C., Martinez Lavanchy, P., Sesink, L., Olivier, B. G., Meakin, J., de Jong, M., & Cruz, M. (2023). *Practical guide to Software Management Plans* (Version 1.1). Netherlands eScience Center and NOW. https://doi.org/10.5281/ZENODO.7589725

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 1–9. https://doi.org/10.1038/s41562-016-0021

National Academies of Sciences, Engineering, and Medicine (NASEM). (2018). *Open science by design: Realizing a vision for 21st century research*. https://doi.org/10.17226/25116

National Academies of Sciences, Engineering, and Medicine (NASEM). (2019). *Reproducibility and replicability in science*. https://doi.org/10.17226/25303

National Institutes of Health. (2014). *Principles and guidelines for reporting preclinical research*. https://grants.nih.gov/policy/reproducibility/principles-guidelines-reporting-preclinical-research.htm

Neylon, C., Belsø, R., Bijsterbosch, M., Cordewener, B., Foncel, J., Friesike, S., Fyfe, A., Jacobs, N., Katerbow, M., Laakso, M., & Sesink, L. (2019). *Open Scholarship and the need for collective action*. Zenodo. https://doi.org/10.5281/zenodo.3454688

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck S., Chambers, C., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D., Hesse, A., Humphreys, M., … Boycan, E. (2023, September 9). *Transparency and Openness Promotion (TOP) Guidelines*. https://osf.io/9f6gx.

Oliver, J. C., Kollen, C., Hickson, B., & Rios, F. (2019). Data science support at the academic library. *Journal of Library Administration*, *59*(3), 241–257. https://doi.org/10.1080/01930826.2019.1583015

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), 943. https://doi.org/10.1126/science.aac4716

Packer, E. (2020, August 24). *ELife launches Executable Research Articles for publishing computationally reproducible results*. eLife. https://elifesciences.org/for-the-press/eb096af1/elife-launches-executable-research-articles-for-publishing-computationally-reproducible-results

Parsons, S., Azevedo, F., Elsherif, M. M., Guay, S., Shahim, O. N., Govaart, G. H., Norris, E., O'Mahony, A., Parker, A. J., Todorovic, A., Pennington, C. R., Garcia-Pelegrin, E., Lazić,A., Robertson, O., Middleton, S. L., Valentini, B., McCuaig, J., Baker, B. J., Collins, E., … & Aczel, B. (2022). A community-sourced glossary of open scholarship terms. *Nature Human Behaviour, 6*(3), 312–318. https://doi.org/10.1038/s41562-021-01269-4

Quan, J. (2021). Toward reproducibility: Academic libraries and open science. In J. Herndon, (Ed.), *Data science in the library: Tools and strategies for supporting data-driven research and instruction* (pp. 49–66). Facet. https://doi.org/10.29085/9781783304615.004

Rahal, R.-M., Hamann, H., Brohmer, H., & Pethig, F. (2022). Sharing the recipe: Reproducibility and replicability in research across disciplines. *Research Ideas and Outcomes*, *8*, Article e89980. https://doi.org/10.3897/rio.8.e89980

Rethlefsen, M. L., Lackey, M. J., & Zhao, S. (2018). Building capacity to encourage research reproducibility and #MakeResearchTrue. *Journal of the Medical Library Association*, *106*(1), 113–119. https://doi.org/10.5195/jmla.2018.273

Ross-Hellauer, T., Klebel, T., Bannach-Brown, A., Horbach, S. P. J. M., Jabeen, H., Manola, N., Metodiev, T., Papageorgiou, H., Reczko, M., Sansone, S.-A., Schneider, J., Tijdink, J., & Vergoulis, T. (2022). TIER2: enhancing trust, integrity and efficiency in research through next-level reproducibility. *Research Ideas and Outcomes*, *8*, Article e98457 https://doi.org/10.3897/rio.8.e98457

Sayre, F., & Riegelman, A. (2019). Replicable services for reproducible research: A model for academic libraries. *College & Research Libraries, 80*(2), 260–272. https://doi.org/10.5860/crl.80.2.260

Stanford Encyclopedia of Philosophy. (2018). *Reproducibility of scientific results*. https://plato.stanford.edu/entries/scientific-reproducibility/

Steeves, V. (2017). Reproducibility librarianship. *Collaborative Librarianship*, *9*(2), 80–89. https://digitalcommons.du.edu/collaborativelibrarianship/vol9/iss2/4

Sumner, J., Haynes, L., Nathan, S., Hudson-Vitale, C., & McIntosh, L. D. (2020). *Reproducibility and reporting practices in COVID-19 preprint manuscripts*. medRxiv. https://doi.org/10.1101/2020.03.24.20042796

The Turing Way Community. (2022). *The Turing Way: A handbook for reproducible, ethical and collaborative research* (1.0.2). Zenodo. https://doi.org/10.5281/zenodo.7625728

Tsang, E., & Maciocci, G. (2020, August 24). *Welcome to a new ERA of reproducible publishing*. ELife. https://elifesciences.org/labs/dc5acbde/welcome-to-a-new-era-of-reproducible-publishing

Wellcome Trust (2017, July 10). *Data, software and materials management and sharing policy.* https://wellcome.org/grant-funding/guidance/data-software-materials-management-and-sharing-policy

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), Article 160018. https://doi.org/10.1038/sdata.2016.18

Notes

---

[1] The "reproducibility crisis", also known as replication or replicability crisis, refers to the observation that a large proportion of scientific studies published across disciplines do not replicate (Open Science Collaboration, 2015). Cf. Framework for Open and Reproducible Research Training (FORRT) Glossary: Reproducibility crisis, https://forrt.org/glossary/reproducibility-crisis-aka-replicab/; Parsons et al., 2022.

[2] Knowledge Exchange: Open Access, https://www.knowledge-exchange.info/projects/project/open-access.

[3] National Institutes of Health (NIH), https://www.nih.gov/research-training/rigor-reproducibility.

[4] Project TIER Protocol 4.0, https://www.projecttier.org/tier-protocol/protocol-4-0.

[5] The Turing Way community and handbook, https://www.turing.ac.uk/research/research-projects/turing-way.

[6] Research Compendium, https://research-compendium.science/.

[7] Knowledge Exchange: Open Access, https://www.knowledge-exchange.info/projects/project/open-access.

[8] New York University: Research Data Management, https://guides.nyu.edu/data_management; Oregon Health & Science University: Research Data and Reproducibility, https://libguides.ohsu.edu/research-data-services/OHSU/Library.

[9] ReproHack Hub, https://www.reprohack.org/.

[10] ReproducibiliTea, https://reproducibilitea.org/.

[11] The Carpentries, https://carpentries.org/.

[12] UKRN: International Reproducibility Networks, https://www.ukrn.org/international-networks/.

[13] E.g. Cascad, https://www.cascad.tech/, Docker, https://www.docker.com/resources/what-container/, GitHub https://github.com/, GitLab, https://about.gitlab.com/, Binder https://mybinder.org/, Whole Tale https://wholetale.org and ReproZip, https://www.reprozip.org.

[14] Project Jupyter community, https://jupyter.org/community; ROpenSci, https://ropensci.org/.