Summary of first TIER2 publisher workshop

AUTHORS:

SUSANNA-ASSUNTA SANSONE ALLYSON LISTER

THOMAS KLEBEL TONY ROSS-HELLAUER

DOI: 10.17605/OSF.IO/TGUXZ

Funded by the European Union

TABLE OF CONTENTS

DOI: 10.17605/OSF.IO/TGUXZ

1. INTRODUCTION

TIER2 is a three-year project funded by the EC and UKRI to boost knowledge on reproducibility, create tools, engage communities, implement interventions and policy across different contexts to increase re-use and overall quality of research results. TIER2 aims to engage and address researchers, funders, and publishers, with co-creation activities central to the project.

On 31st of May 2023, TIER2 convened its first workshop for publishers, attended by 20 representatives of major publishers. The meeting agenda and list of attendees is provided in the Appendix.

The workshop's aim was to:

- Introduce the TIER2 project.
- Discuss initiatives, measures, and interventions, in development or in place, to increase reproducibility of published research.
- Identify and prioritise new or existing areas for further development, where TIER2 could either contribute to, or drive, developments.

In advance of the workshop, participants completed a short questionnaire to capture their experiences regarding the main challenges. The compiled findings were then presented at the workshop and used as the basis for discussion.

This report summarizes the outcomes of the survey and workshop outcomes. In addition, we present TIER2's ideas for two co-creation activities that we have conceived based on these discussions.

2. SURVEY OUTCOMES

QUESTION 1

"What are your main challenges to increase or assure the reproducibility of the research you publish?"

DOI: 10.17605/OSF.IO/TGUXZ

All 20 participants answered this question, usually identifying multiple issues within the same answer. The answers are clustered under the following themes and ordered by number of responses:

Capacity, costs, and demand

Lack of time and/or resources on the publisher side, as well as a desire to avoid further burden on authors, reviewers, and academic editors or increasing turnaround times for publications. Issues of scalability and the need to build upon existing infrastructures (with subsequent costs and limitations) were also mentioned, along with a lack of current demand from authors for such services.

Data sharing practices and support

Promoting and implementing sharing of data (including datasets, code, protocols, materials, and methods) underpinning publications are still challenges. In some cases, there are also technical issues in (assisting with) hosting very large datasets, checking the quality of such outputs, and in enforcing journal data policy. In particular, Data Availability Statements (DAS) are perceived to be ineffective when they allow statements such as "data available upon request."

Incentives

Lack of incentives for all players, including researchers, to embrace open and reproducible practices, and for publishers and editors and reviewers to take extra steps to ensure best practices.

Knowledge

Lack of education and awareness. This ranged from a lack of basic understanding of reproducibility, with highly variable levels of knowledge amongst researchers (at all career levels) noted, to the need for more training for all parties involved (publishers, editors, reviewers, and authors).

Peer review and editorial checks

Challenges related to checks on adherence to community reporting standards, the FAIR Principles, and enabling (inter-)linkages among articles and related code, protocols and preprints, and other relevant digital objects. In addition, basic difficulties of reviewer recruitment were noted as a perennial issue also affecting any kind of checks.

Other stakeholders

A concerted mobilisation of all stakeholders in the scholarly ecosystem, including institutions and funders, is essential to reform incentives, educate, and define policies to enable researchers to engage with good data practices. Others, such as tool and services developers, are also pivotal in implementing new workflows and resources (e.g., Registered Reports) to support the reproducibility of published research.

One size does not fit all

Difficulties in finding "efficient technical solutions" that work across journals and multiple disciplinary areas, especially considering differences in the relevance of reproducibility across research practices, as well as differences in disciplinary norms.

Bad actors

These include paper mills, peer review/citation cartels, and undeclared usage of Large Language Models, although the growing visibility of such issues potentially present an opportunity to raise awareness of reproducibility issues and hence expedite progress.

QUESTION 2

"What are you already doing to support reproducibility?"

The answers are clustered under the following categories and ordered by number of responses:

Sharing of data and other digital objects

Actions enacted range from strict mandates (e.g., requirements for data/code deposition) to looser policies (e.g., data availability statements "a recent audit shows this is ineffective"), to promotion of other measures such as protocol sharing.

Community reporting guidelines

Promotion of community reporting requirements. In-house checks are done by some journals for adherence to a very selected number of checklists in the clinical/biomedical space (e.g., CONSORT, PRISMA, STROBE, CARE, ARRIVE, MDAR), and authors are encouraged to make statements regarding constraints on generality (COG) or how research meets TOP guidelines.

DOI: 10.17605/OSF.IO/TGUXZ

Peer review and editorial workflows

Improved support for checks on availability and quality of data and code. One respondent indicated that some journals in their portfolio had taken on data or reproducibility editors as part of the academic editorial team, with the aim of increasing data availability and checks on data and code. Others mentioned asking reviewers specifically to look at data and code, engaging with data-focussed reviewers, and curators assisting in FAIRification of data, as well as supporting open peer review to increase transparency of quality assurance processes.

Publication types

Supporting a broader range of publication types to increase reproducibility, including registered reports, null results, reanalysis articles, protocols, data notes, software tools, methods articles, and replication studies as new article types.

Partnerships

Pursuing collaborations with preprint servers, protocol databases (e.g., protocols.io) and code/data repositories (e.g., Code Ocean) to facilitate sharing and interlinking of digital objects.

Other respondents mentioned supporting Open Science badges, open metadata, not promoting impact factors and meta-research to understand researchers' practices and needs as other enabling measures already in place.

QUESTION 3

"Which are the key issues or initiatives to boost reproducibility of published research that you would like to see realized?"

16 responses, each listing one or more issue, were clustered under the following categories:

Reform of incentives

7 respondents mentioned the urgent need to change the incentive system, including credit for full sharing of data/methodologies, rewards for work that proves reproducible (as opposed to "flashy") results, and incentives to conduct and publish replication studies. However, as highlighted in answers to Q1, this requires a concerted mobilisation of all stakeholders in the scholarly ecosystem, including institutions and funders.

DOI: 10.17605/OSF.IO/TGUXZ

Stronger policies

7 respondents wished for firmer requirements for enhanced reporting and sharing of data and other digital objects. Also mentioned were better reporting of processes of data generation and analysis and increased pre-registration of studies.

Joined-up approaches

6 respondents saw greater cross-sector collaboration as a key future priority, on elements like agreement on priorities and multi-stakeholder strategy, agreed terminologies for elements of reproducibility, and ways to improve uptake of key practices.

Standards

5 respondents identified development of standards as a key priority, including for structured FAIR data standards (e.g., packages), enhanced interlinkage of research objects, uptake, and development of the TOP guidelines, and (in the view of one) development of a body (similar in structure to COPE) that could coordinate discussion on general common standards for publishers.

Review and editorial checks

3 respondents foresaw enhanced review and editorial checking as important, including checks on adherence to community reporting standards (with automation where possible), properly resourced quality control and data curation services, and transparent peer review.

Training

3 respondents mentioned the need for improved training and education on Open Science practices for researchers/authors, including the benefits of sharing. One mentioned, however, that greater attention should be given to overcoming differences in levels of knowledge depending on region.

Monitoring

Finally, one respondent wished for better measurement and monitoring of transparency and reproducibility, and their impact on research quality and reuse.

DOI: 10.17605/OSF.IO/TGUXZ

3. PRESENTATIONS

To provide additional context to the discussion, six participants contributed lightning talks to elaborate on some of the community initiatives they are involved in, or relevant activities they carry out in house.

Bernd Pulverer Head and Chief Editor at EMBO Press	Presented a range of initiatives that EMBO are currently engaged in, including PRO-MaP, work with curators to ensure adherence to source data mandates, and the CREC NISO working group to improve the effectiveness and visibility of corrections to the published literature.
Chris Hunter GigaDB director at GigaScience Press	Presented workflows for conducting audits of data availability for GigaScience outputs via checklists and annotation tools.
Elisa De Ranieri Head of Research Integrity and Author Experience at Cell Press	Presented PRO-MaP (Promoting Reusable and Open Methods and Protocols), newly published recommendations for researchers, research institutions and departments, publishers and editors and funders to improve the reporting of detailed, reusable and open methods and protocols. While developed within the life sciences, the recommendations are intended to be applicable to other disciplines.
Guy Jones Chief Editor at Springer Nature	Presented initiatives to streamline policies on data sharing at Springer Nature. Whereas previously there were four policies across the portfolio, now there is just a single "baseline" policy: "Data Availability Statements are required for all primary research articles."
Iain Hrynaszkiewicz Director, Open Research Solutions at PLOS	Presented PLOS' work to create indicators for Open Science. Noting the need for improved ways of measuring uptake of key Open Science practices, PLOS has been partnering with Dataseer on approaches to quantifying levels of preprinting, data-, and code-sharing within PLOS articles (and benchmarking against similar journals).
Matt Cannon Head of Open Research at Taylor & Francis	Presented a T&F initiative to improve adherence to the TOP Guidelines by piloting optional author statements within published manuscripts on how authors have met the eight TOP guideline criteria.

The slides for the presentations are available here.

DOI: 10.17605/OSF.IO/TGUXZ

4. DISCUSSION POINTS

The following points were also raised during the workshop discussion:

- Reform of incentives is seen as central to enabling real change and encouraging uptake of practices such as Registered Reports and publication of replication studies. However, the realignment of incentives requires concerted effort from all stakeholder groups to ensure complementarity and alignment.
- The role of metrics is double-edged. On the one hand, making practices associated with reproducibility more measurable and visible could incentivize uptake, on the other there is always the risk that the measure will become the target, with potential negative implications.
- Joined-up approaches yield stronger results. At the level of publishers, such approaches could overcome some of the issues about resourcing and the tension between the “school of efficiency and school of integrity”. Interventions are more likely to work if adopted by many journals, using common standards; if multiple journals move together, authors will also be incentivized to engage. Also, common 'myth busting‘ material could help to inform and educate authors around, for example, concerns about making data available.
- Definitions of reproducibility across fields: Different kinds of reproducibility might be desired/possible depending on the field or mode of knowledge production which makes policy harmonization across publisher portfolios difficult. In addition, the range of definitions for various aspects of reproducibility, with confusion over terms, was seen as a barrier.
- Vulnerabilities should be protected and imposing Global North standards elsewhere will further increase inequity without financial support for infrastructure and capacity building. There are concerns about disadvantages among authors when sharing data, and valid reasons not to share must be considered in reform of policies and incentives (e.g., protected data badge).

5. NEXT STEPS

There was discussion on where TIER2 could make the most impact upon these issues within its run-time. Although reform of incentives was seen as the area with the potential to make the largest impact, there was acknowledgement that such structural changes are a long-term issue where change might not be achievable within TIER2. The resources required to achieve impact in some areas, including strengthening policies, training, and monitoring, were also noted.

DOI: 10.17605/OSF.IO/TGUXZ

Following the workshop, in follow-up discussions, we have conceived two pilot activities which are designed to achieve maximum impact within our project.

5.1 PILOT HANDBOOK

While some journals have internal guidance on promoting and enabling reproducible and FAIR data, there is little/no consensus among publishers. The planned co-creation and test of an educational reference handbook will help operationalize data checks to assist reproducibility and provide editors with a harmonized set of data checks. This handbook can also serve as advice to authors and reviewers and contribute towards reproducibility and FAIRness.

TITLE

Reproducibility Handbook for editors and authors

AIMS

The main objective is the co-creation and testing (via intervention) with interested publishers of training material for editors. This handbook will be both educational and operational guidance in support of reproducibility and FAIRnesss.

OUTCOME

This work will create a reproducibility handbook containing a core set of reproducibility checks (with definition, value, examples, and implementations) and training/educational material covering common steps of these checks; such a shared handbook will serve editors and reviewers, as well as make these internal processes transparent and understandable to the authors.

TIMELINE

Preparation and planning (Dec 2023 – Feb 2024)

- Organisational meetings and first publisher meeting to define the core reproducibility checks.

Development and feedback (Mar 2024 – Oct 2024)

- Multiple rounds of workshops, collection of internal feedback by publishers, revisions on core checks

Implementation and Evaluation (Jan 2025 – Dec 2025)

- Active intervention (Jan 2025 – Aug 2025)
- Post-intervention survey among editors
- Data analysis and publication writing

DOI: 10.17605/OSF.IO/TGUXZ

5.2 PILOT DAS

An intervention around Data availability statements (DAS), improving and then piloting their use, by working with the RDA Policy group, also driven by the findings of the STM Research Data's survey. TIER2 can both contribute and help with organizing co-creation events.

TITLE

Editorial workflows to increase data sharing

AIMS

Data availability statements (DAS) related to the data underlying publications are currently not effective at promoting data sharing and reproducibility. Editors have substantial leverage to request clarifications from authors on data sharing. The planned intervention will provide editors with straightforward workflows to improve DASs, for example by pushing back on problematic statements such as “data available on request”.

METHODS

The intervention will be developed via co-creation, conducting workshops with publisher representatives and editors.

OUTCOME

(1) An editorial workflow that improves data sharing via feedback on Data availability statements.
(2) Empirical evidence on the efficacy of the workflow across multiple journals/contexts.

TIMELINE

Planning and development of the workflow (Jan 2024 - May 2024)

- Workshops with publisher representatives and editors

Implementation of pilot intervention at participating journals (June/July 2024)

- Sample selection and preregistration of intervention methodology

Intervention active (July 2024 – March 2025)

Analysis and write-up (April 2025 – Sept 2025)

- Post-intervention survey among editors
- Data analysis and publication writing

DOI: 10.17605/OSF.IO/TGUXZ

6. APPENDIX – AGENDA AND ATTENDEE LIST

6.1 AGENDA AND SLIDES

(all times CEST)

16.00 - 16.20	**Welcome and introduction (Tony)**
16.20 - 17.00	**Overview of survey responses (slides)** ● Main challenges for ensuring reproducibility in publishing (Liz) ● Current initiatives (Tony) ● Main priorities (Susanna) **Discussion**
17.00 - 17.20	**Stakeholder presentations (6 X 3 mins) (slides)** ● Bernd Pulverer, EMBO ● Chris Hunter, GigaScience Press ● Elisa De Ranieri, Cell Press ● Guy Jones, Springer Nature ● Iain Hrynaszkiewicz, PLOS ● Matt Cannon, Taylor & Francis
17.20 - 18:00	**Discussion and prioritisation of new solutions**

6.2 ATTENDEE LIST

Organisers	Affiliation	Role in organisation
Tony Ross-Hellauer	TU Graz	Group Leader, and TIER2 PI
Liz Allen	F1000	Director of Strategic Initiatives
Susanna-Assunta Sansone	University of Oxford, and FAIRsharing	University Academic Lead of Research Practice; Prof of Data Readiness, Dep. of Engineering

DOI: 10.17605/OSF.IO/TGUXZ

Invitees	Affiliation	Role in organisation
Andrew L. Hufton	Patterns	Editor-in-Chief
Anita Dewaard	Elsevier	VP Research Collaborations
Annie Hill	American Psychological Association	Editorial Director, Journals
Bernd Pulverer	EMBO Press	Head and Chief Editor
Catriona J. MacCallum	Hindawi	Director of Open Science
Chris Hunter	GigaScience Press	GigaDB director
Elisa De Ranieri	Cell Press	Head of Research Integrity and Author Experience
Guy Jones	Springer Nature	Chief Editor
Iain Hrynaszkiewicz	PLOS	Director, Open Research Solutions
Kiera McNeice	Cambridge University Press	Research Data Manager
Matt Cannon	Taylor & Francis	Head of Open Research
Michael Streeter	Wiley	Director, Research Integrity & Publishing Ethics
Molly Cranston	F1000 (Taylor & Francis)	Editorial Content Manager

Invitees	Affiliation	Role in organisation
Nick Lindsay	The MIT Press	Director for Journals and Open Access
Nicola Nugent	Royal Society of Chemistry	Publishing Manager, Quality & Ethics
Phil Hurst	The Royal Society	Publisher
Pierre Nauleau	Lancet	Senior editor
Rhodri Jackson	Oxford University Press	Publishing Director
Theo Bloom	BMJ	Executive Editor
Wei Mun Chan	eLife	Editorial Manager and Research Integrity Advisor