



CHORUS Forum: Research Data Management

Research Data Policies and the Publishing workflow

Dr. Rebecca Taylor-Grant, Director of Open Science Strategy & Innovation, Taylor & Francis

20 May 2026

Publishers as a data sharing stakeholder

- Policies can be enforced at a key point in the research lifecycle
- Journal policies are a common point for other stakeholder data policies may be enforced (e.g. CRUK, Wellcome Trust, NIH)
- Key outcome is the use of Data Availability Statements to explicitly describe how datasets can be accessed

Author contributions

All authors have read and approved the final work.

Conceptualization: FC, MPC, KA, SD, NV; Methodology: NV, KA, DR, SL, ML, FG, MG; Validation: MPC, SD, KA, DR, NV, FG; Formal analysis: NV, ML, DR; Investigation: NV, MG, AF; Resources: MPC, SL, FG, DR; Data curation: NV, ML, MG; Writing – original draft preparation: NV, ML, MPC, SD, DR, FG; Writing – review and editing: NV, ML, MPC, SD, DR, FG; Visualization: NV, ML, DR, MG; Supervision: MPC, FC, SD, KA; Project administration: MPC, SD, FC, KA; Funding acquisition: MPC, FC

Data availability statement

The data that support the findings of this study are openly available in figshare (<https://figshare.com>) at <https://doi.org/10.6084/m9.figshare.25651701>, reference number 25,651,701. The raw sequencing data and assembled genomes used in this study have been deposited in the Sequence Read Archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>) under NCBI BioProject accession no. PRJNA900024. SRA and assembly accession numbers are listed in Table S2 in the supplemental material.

Preprint publication

A preprint version of this manuscript was submitted to bioRxiv and is available at the following link <https://doi.org/>

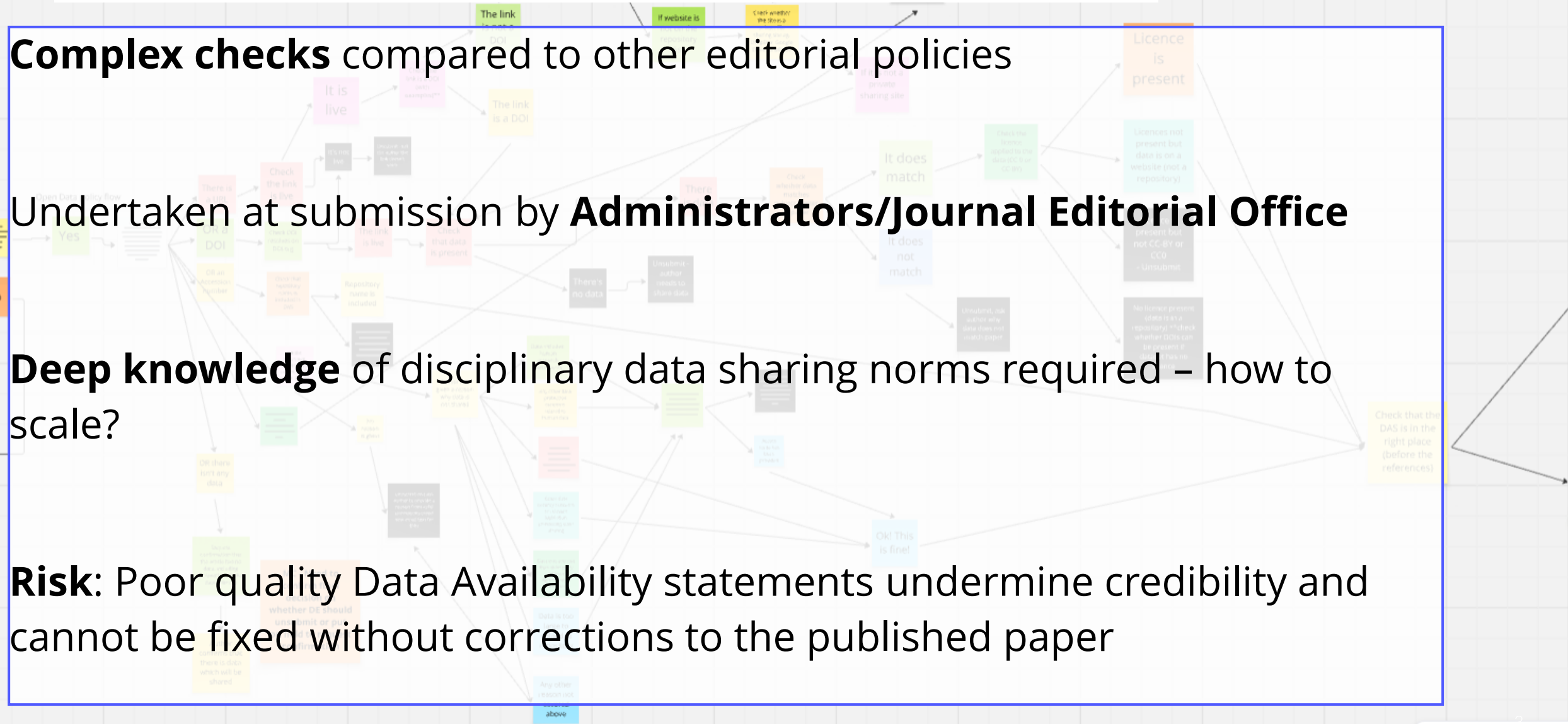
The problem with data policy enforcement

Complex checks compared to other editorial policies

Undertaken at submission by **Administrators/Journal Editorial Office**

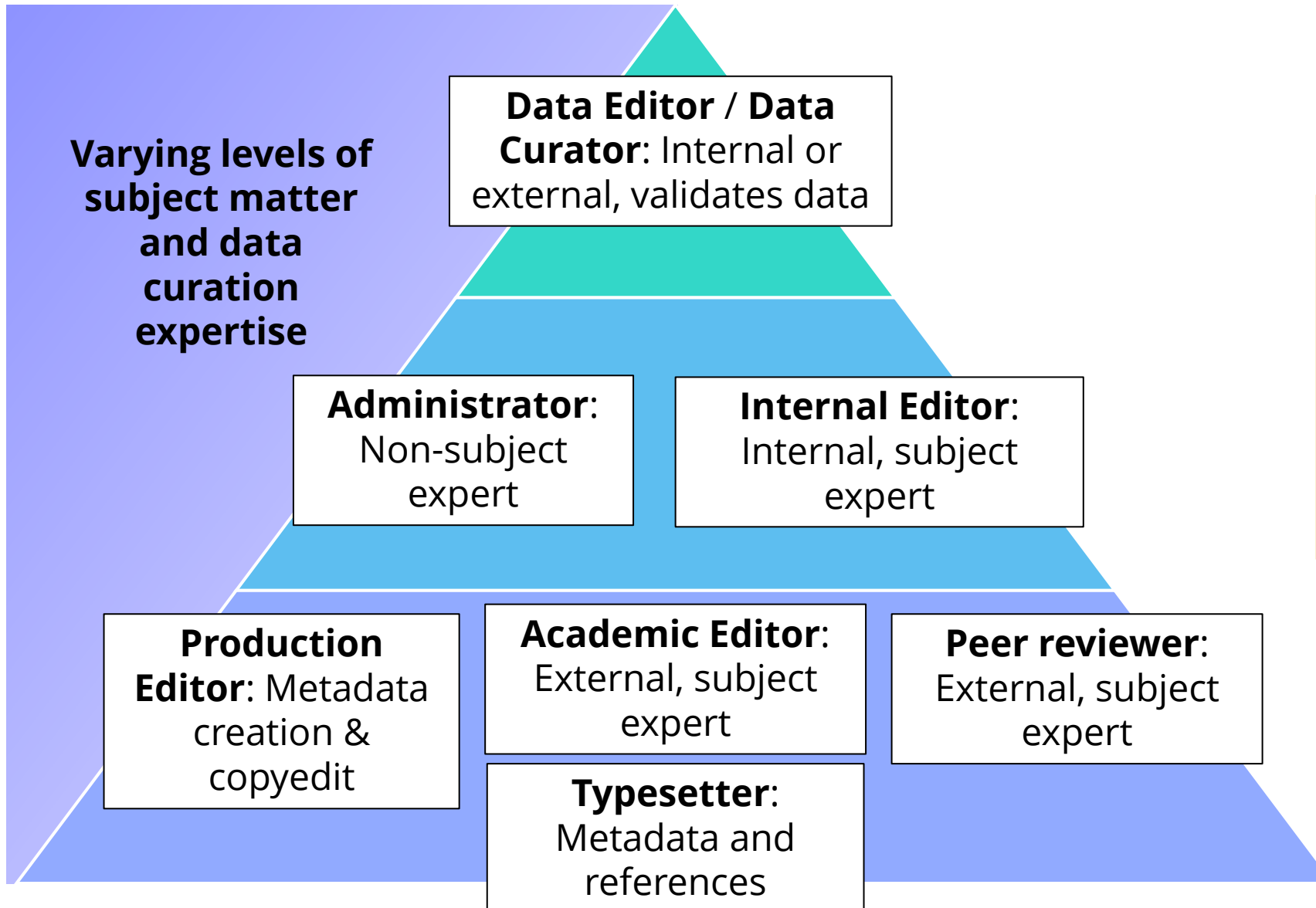
Deep knowledge of disciplinary data sharing norms required – how to scale?

Risk: Poor quality Data Availability statements undermine credibility and cannot be fixed without corrections to the published paper



Supporting better manuscript checking: the TIER2 project

Data policy and manuscript handling



Problem of scaling: data curators and editors are difficult to recruit for, and data curation expertise of academic editors and peer reviewers varies



Editorial Reference Handbook

Overview

Version of 16 July 2024; CC BY-SA 4.0.

The Editorial Reference Handbook (hereafter the 'Handbook') contributes towards a common understanding of what is required by scholarly journals and publishers to assist reproducibility and FAIRness in practice.

Rationale and aims

Components of the Handbook

Start using the Handbook

 Export as PDF

Developed collaboratively with publishers to support the publication of reproducible and FAIR research

Including:

- ✓ Guidance for editors on implementing more stringent data checks (e.g. enforcing stronger data sharing policies).
- ✓ A list of checks.
- ✓ A model workflow with roles.

The Handbook:

➤ **Outlines the fundamental requirements to enforce data sharing policies**

➤ **Provides an overview of what is needed to aid decision making**

➤ **Identifies key roles and responsibilities**

<http://www.doi.org/10.17605/OSF.IO/KNQZU>



Taylor & Francis
by informa•••

Using AI to further enhance data checking workflows: DataSeer SnapShot

Common errors made by authors in submitted manuscripts



Data availability statements are often **missing**



Data “**shared on request**” when open data sharing is required



Data shared via **private file sharing links** instead of repositories (e.g. Dropbox, Google Drive)

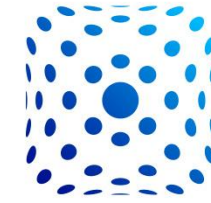


URLs or peer review links shared instead of PIDs/DOIs

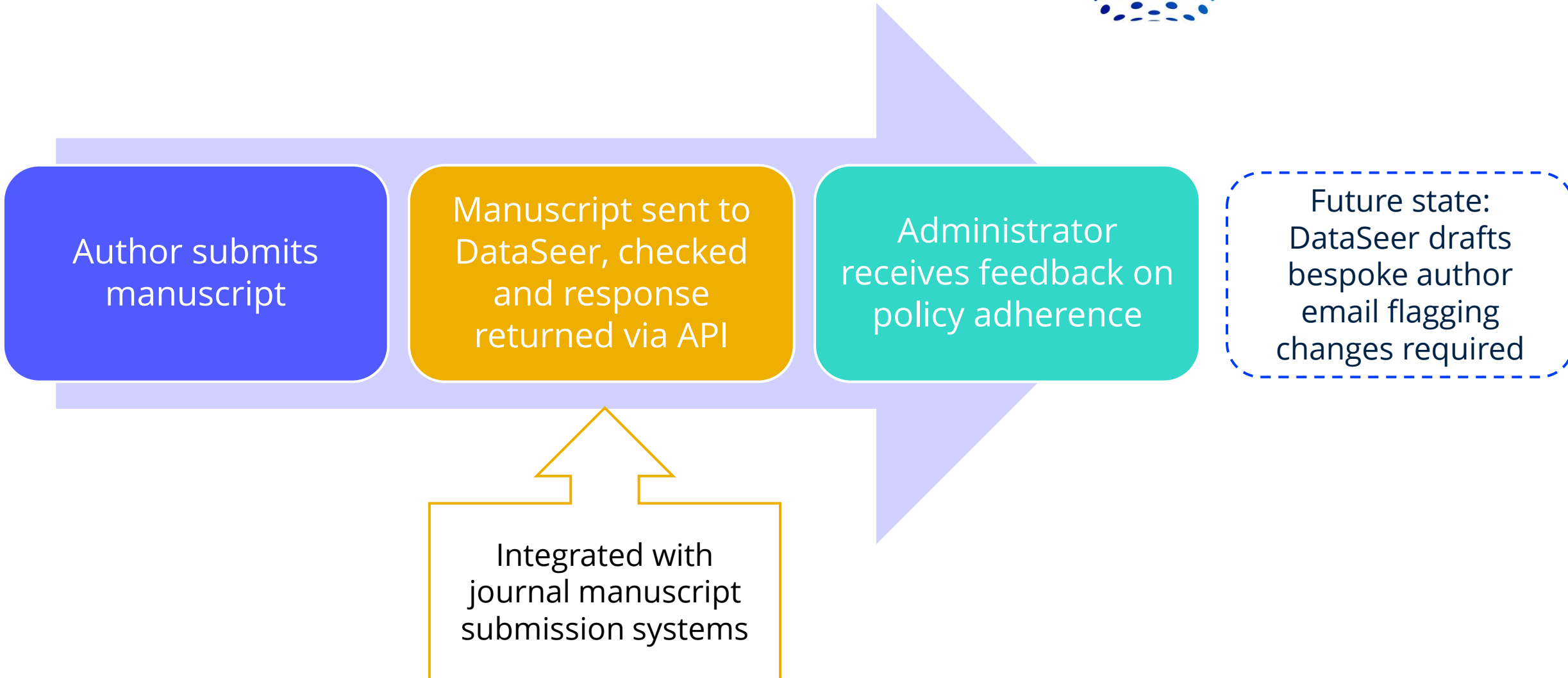


Unacceptable exemptions requested (e.g. data not shared as the author is still using it)

Embedding data checking in the editorial workflow



DataSeer



PCOMPBIOL-D-24-01303-R1-Editor-Author_Version.pdf


Sparse high-dimensional decomposition of non-primary auditory cortical receptive fields

QUESTIONS ABOUT THESE RESULTS?

LLM CHAT

Action Summary








Data are on acceptable repository.

 No Action Required

Data Availability

ITEM	STATUS	NOTES / ACTION
Editorial Manager DAS	 Present	Data and software implementations in MATLAB v2017b of the algorithms discussed are available at: https://hdl.handle.net/1903/33143
Manuscript DAS	 Absent	

Checklist

ITEM	RESPONSE	NOTES / ACTION
Does the DAS state "All data are in the manuscript and/or SI files"?	 No	
Will the authors submit their data to a repository upon acceptance?	 No	
Check for the minimal dataset in the SI files.	 N/A	
Does the DAS mention restrictions to data availability?	 No	
If access restrictions apply, is an author the contact point for data requests?	 N/A	
Does the repository meets PLOS's standards?	 Yes	
Are there any non-functional URLs?	 No	

In summary



Introducing a data policy and enforcing a data policy are not the same

Introduction of data policies can introduce significant complexity into the manuscript checking workflow

We have not always had standardised recommendations on policy enforcement – this is changing

Significant potential of AI to reduce administrative burden in future