Joint Position Statement on "Data Repository Selection - Criteria That Matter"

Issued by Confederation of Open Access Repositories, CoreTrustSeal, European University Association, Science Europe, and World Data System

April 1, 2021

Over the past three years, "Data Repository Selection-Criteria That Matter"[i] – "a set of criteria for the identification and selection of those data repositories that accept research data submissions" – were developed by a group of publishers facilitated by the FAIRsharing initiative. Throughout this time, a large number of organizations and individuals have formulated responses[ii] and expressed concern[iii] about the criteria and the process through which the criteria were developed.

Community Concerns

This activity has alerted different stakeholders in the research community for a variety of reasons:

- It will cause publishers to exert undue influence on researchers' decisions about the appropriate curation and preservation care needed for their data

- It threatens to exclude many repositories and limit options for researchers

- It does not build on existing good practice frameworks for data repositories

- It is not driven by public interest and it will create an unreasonable dependency on a single repository registry which does not include researchers and repositories in its governance

- It could conflict with funder, institutional or national policies/legislation. In particular when researchers are mandated to deposit data in specific repositories by their institutions, funders or legislation

- And ultimately, it will be detrimental to the advancement of open science across the world

Although the development process included a request for input, many in the research data community have felt their input has not been incorporated into subsequent versions of the paper, and the only response from the authors did not adequately address these concerns[iv].

We are issuing this joint position statement to highlight the community's concerns and request that the group respond with the concrete actions outlined below.

Good Practices vs. Features

We maintain that data repositories should support the work of researchers with the highest quality standards regardless of the scientific discipline. Registries should exist to help researchers make a choice about the appropriate repository for them, based on their needs, not the desires of a special interest group.

The "Criteria that Matter" approach conflates the notion of "good practice criteria" with repository "features". Repository criteria define practices developed and agreed on by the community, which address specific objectives, such as preservation, sustainability, accessibility, and re-use (e.g. trustworthiness, FAIR data) and are used to help repositories improve their practices in managing research data. Features, such as badges, publication linking mechanisms, or data access for pre-publication review, are functionalities that may be desired by a given stakeholder community to support their specific needs or use cases. As currently presented, the "Criteria That Matter" repeats some good practice criteria that are already covered by existing and widely adopted community frameworks, as well as requiring a limited set of features that matter primarily to publishers, some of which could conflict with funder, national or institutional policies.

The existing good practice frameworks for data repositories and data management were developed through widespread community input and vetting and are aligned with "The TRUST Principles for Digital Repositories", the "FAIR Data Principles", and the "CARE Principles for Indigenous Data Governance". These principles, expressed via consensus, summarise the expectations and needs of the community at a high level. They have been or are being translated into frameworks of good practice, benchmarks and/or criteria for assessing compliance such as the "CoreTrustSeal requirements for trustworthy data repositories" or the "Science Europe Practical Guide to International Alignment of Research Data Management". The practical implications of adopting these community-developed criteria and aligning data repositories' practices and capabilities to support common objectives are being explored by different stakeholders.[v] This includes efforts to ensure that these frameworks do not exist in isolation, but are connected in meaningful ways.[vi]

The selection of repositories should be primarily guided by the needs of the research community and aligned with the criteria already covered in these community-developed frameworks. While it is legitimate for publishers to engage with repositories to define the features that matter to them and expose them in registries, they should not be deciding what

repository functionalities are important for the research community and such features should not be presented as criteria for the selection of appropriate repositories.

Publishers' guidance to support researchers selecting a repository should be aligned with, and not substitute or conflict with, guidance already available to researchers from their institutions, disciplinary communities, and/ or funders.

We are requesting the authors to:

1. Take into consideration the concerns expressed by the community and reconsider their current approach to focus their contribution on where it can have the most value for the research community

2. Not restrict researchers' options or exclude repositories unless based on relevant research community needs and scientific criteria

3. Engage in a transparent manner with the community including existing recognised community efforts to define best practices for repositories

4. Document and define which features publishers consider a priority for the data that underpins publications, clearly justify any publisher-specific features, and engage with repositories about implementing them

5. Include representatives of the major communities (funders, researchers, institutions, and repositories) in the development and governance of publisher-specific features

6. Any implementation of the publisher-specific features in repository registries, including FAIRsharing, should be contingent on the above. Changes to the repository description metadata and repository filtering functionality implemented in registries should be transparent and aligned with an ongoing consultation process

••

This position statement is also supported by a number of other organizations listed below.

Supporting Organizations

- AGU: American Geophysical Union
- Canadian Association of Research Libraries and Portage Network
- CSIC: Spanish National Research Council
- COPDESS: Coalition for Publishing Data in the Earth and Space Sciences
- EIFL: Electronic Information for Libraries
- LA Referencia
- LIBER: Association of European Research Libraries
- LIBSENSE Africa
- OpenAIRE
- National Committee for Data in Science of the Australian Academy of Science
- SPARC: Scholarly Publishing and Academic Resources Coalition
- SPARC Europe

If you would like to support this Joint Position Statement or request more information, please contact office@coar-repositories.org

---

References

[i] https://doi.org/10.17605/OSF.IO/N9QJ7

[ii] For example, FAIRsFAIR https://doi.org/10.5281/zenodo.4630588 and CoreTrustSeal https://doi.org/10.5281/zenodo.3632916

[iii] https://www.coar-repositories.org/news-updates/input-to-data-repository-selection-criteria-that-matter/

[iv] https://doi.org/10.5281/zenodo.4458082

[v] Frameworks have been developed to enable other objectives building on TRUST and FAIR principles: COAR Framework for Good Practices in Repositories, Science Europe Practical Guide to the International Alignment of Research Data Management, World Data System's (WDS) Members accreditation, The American Geophysical Union's (AGU) Guidelines to support author compliance with open data standards, The US Office of Science and Technology Policy (OSTP) Desirable Characteristics of Repositories for Managing and Sharing Data Resulting From Federally Funded Research, The Common Language Resources and Technology Infrastructure (CLARIN) ERIC (European Research Infrastructure) B centres Assessment procedure.

[vi] FAIRsFAIR project work on aligning trustworthiness with FAIR and the FAIRification of repositories. See, for example Hervé L'Hours, Ilona von Stein, Frans Huigen, Anusuriya Devaraju, Mustapha Mokrane, Joy Davidson,Robert Huber. (2020, August 31). CoreTrustSeal plus FAIR Overview (Version 03.00). Zenodo. http://doi.org/10.5281/zenodo.4003630.