

In response to the:

COAR Blog <https://www.coar-repositories.org/news-updates/input-to-data-repository-selection-criteria-that-matter/>

OpenAIRE Blog <https://www.openaire.eu/openaire-in-support-of-coars-input-to-data-repository-selection-criteria-that-matter>

and the LIBSENSE community.

We thank both the COAR, OpenAIRE and LIBSENSE communities for their feedback and comments on our preprint. Firstly, we (the below signatories) agree with the aim stated in the COAR blog post to, 'develop a...network of repositories that can support researchers around the world in managing and sharing their data'. Publishers are one of the user groups of the repository network, and as such we are glad to collaborate with our fellow stakeholders to facilitate the shared aim of supporting research data sharing globally.

The motivation for publishers in facilitating the use of research data repositories is to fulfil our core responsibility of maintaining the integrity of the scholarly record. Researchers as authors often approach publishers requesting information about how best to share the data that accompanies their research articles. We recognised there was duplication of effort among publishers in trying to advise authors on best practices, and occasionally advice would differ from publisher to publisher. Rather than looking to mandate authors to host their data in specific repositories or recommend different repositories over others, the immediate aim of this project has been for publishers to work together to understand the information we are providing authors about data sharing in repositories. The ultimate aim is to increase the amount of research data stored in the most suitable repository. We see it as the author and their research community's decision as to the optimal location for research data and institutional or regional repositories are suitable options for many researchers.

Many of our proposed features match existing frameworks from CTS, TRUST, COAR, NIH, and ELIXIR (see our comparison [table](#)). The features we describe in our preprint specify information we consider useful for decisions to be made about the best venue for research data. These fields are to encourage transparency of this information, rather than to be used as 'criteria' for 'compliance' and include fields useful to publishers in determining how best to support authors in meeting data sharing requirements. Two notable examples are:

- "Data Access for Pre-Publication Review"
This is useful for authors who want to make data available in a repository for review prior to article publication. However, authors may ultimately archive their data to a repository which does not have this feature.
- "Data Deposition Condition"
This is useful for publishers to know whether or not there are limitations on who can deposit data in a repository; for example, some repositories only allow data deposits from members of a university, whilst other limit data deposition to national researchers. This information makes it possible to provide appropriate repository recommendations that would be open to all authors.

The publishers involved in the project already offer resources to assist authors who need help with repository selection, such as lists of recommended repositories. However, currently the reasons why certain repositories are recommended is not always transparent. Some major publishers require repositories to apply to be included in lists. Other lists have been created to provide assistance to authors who ask for it and are not often updated. One of the objectives of this project is to increase transparency about why publishers recommend specific

repositories, and the features they encourage researchers to look for when choosing a repository.

With a set of features like those proposed, publishers should be able to offer authors more consistent, transparent advice that will help researchers weigh up features available from different repositories themselves, and to make well-informed choices.

We are thankful for the comments received from both COAR, OpenAIRE and LIBSENSE, to add to those received via a survey of 53 stakeholders -- including repository owners, major data services and organizations. This has already resulted in over 115 comments. We continue to invite and accept community feedback as we refine and improve this work.

Based on all this feedback will be revising the preprint (<https://zenodo.org/record/4084763#.X9dKVtj7Q2w>) ahead of submission to a journal for formal publication. To more closely reflect the sentiment behind the project, which we have outlined in this response, we are considering some alternate wording. For example, referring to 'features' of the repositories, rather than 'criteria'. This should clarify that what we are presenting here is intended to be advice for authors who are looking for it, rather than a directive to influence how and where data is stored. We will also include the table we have created, mapping repositories' features from relevant complementary initiatives, showing that overall there is an agreement and convergence on core elements.

We are open to discussing your feedback in more detail to ensure this is a resource that complements the work of the COAR, OpenAIRE and LIBSENSE communities, to support our mutually held goal to support research data sharing globally.

Signed by the repocriteria@googlegroups.com members:

Matt Cannon	Taylor & Francis
Chris Graf	Wiley
Catriona MacCallum	Hindawi
Jonathan Threlfall & Molly Cranston	F1000
Sarah Callaghan & Ilaria Carnevale	Elsevier
Varsha Khodiyar	Springer Nature
Scott Edmunds	GigaScience Press
Thomas Lemberger	EMBO Press
Wei Mun Chan	eLife
Kiera McNeice	CUP
Iain Hrynaszkiewicz	PLOS
Adam Leary	OUP
Peter McQuilton & Susanna-A Sansone	FAIRsharing