Australian Research Council Consultation: Inclusion of preprints in the Excellence in Research for Australia data collection

21 March 2022

Science & Technology Australia (STA) is the peak body representing more than 90,000 scientists and technologists in Australia.

Thank you for the opportunity to offer input to the Australian Research Council's (ARC) consultation on the potential inclusion of preprint publications in the Excellence in Research for Australia (ERA) data collection.

After consultation with our diverse membership, our view is that while preprints are gaining broader acceptance in some disciplines and are an increasingly important way of tracking cutting edge research, it is not appropriate to include preprints in the ERA data collection.

As most preprints progress to fully peer-reviewed published articles, it is this final product that is more appropriate to include in ERA assessments.

This is particularly true given that the ERA reference period spans several years, during which many preprints would progress through to a journal publication.

Additionally, given ERA's focus on measuring quality and research excellence, the absence of a consistent formal peer review process for all preprints renders them not appropriate for inclusion as an eligible research output.

There are some limited instances where publication on a preprint server may be the ultimate publication mechanism and the primary way to disseminate research results and new knowledge for a specific piece of research within some fields. These could be considered for inclusion in measures of total research output/productivity, but very specific criteria to ensure quality (a minimum threshold of peer review) and uniqueness (a DOI or specific URL assigned to the publication) must be issued to uphold the integrity of the ERA data collection. This issue warrants further targeted consultation with representatives from these fields.

1. Given ERA's focus on evaluating research excellence, is it appropriate to consider inclusion of preprints in ERA noting the issues with the indexation of preprints? *Yes/No. Please explain your answer.*

No.

While publishing research results and findings via preprints repositories is now largely accepted in some disciplines, this is not the case across all STEM disciplines – the majority of which still use traditional journals as the dominant medium for publishing research. The majority of preprints will also progress to a published, peer-reviewed journal article. Given ERA's focus on measuring research excellence, this published journal article is the appropriate item to include in the ERA assessment.

For some fields, publication in a preprint repository may be the ultimate publication mechanism and the primary way to disseminate research results and new knowledge. To measure total research output/productivity in these fields, inclusion of preprint publications could be considered – if publication via preprint servers has become the norm for a discipline, this should be reflected in the ERA assessment of that discipline. Inclusion of preprints as an eligible research output for specific 2-digit FoRs could be considered on a trial basis, comparing the results of assessments both with and without preprints included. To maintain the focus on research excellence, there would need to be very specific criteria to clearly identify which publications would be eligible, ensuring quality (a minimum threshold of peer review) and uniqueness (a DOI or specific URL assigned to the publication).

Our members have drawn a distinction between enabling researchers to cite preprints in proposed research under ARC grants schemes – which they strongly support and was the subject of research sector advocacy last year – and having preprints counted in ERA, which they do not support overall.

2. Is it appropriate to include preprints as an eligible research output type for citation analysis disciplines? Yes/No. Please explain your answer.

No.

Broadly speaking, citation analysis assessment should include fully peer-reviewed journal articles only.

As mentioned above, for specific 2-digit FoRs, it might be appropriate to include preprints which are themselves the final publication product – published on a preprint server with a unique DOI or URL, and that have undergone a minimum threshold of peer review. There would need to be very specific criteria to clearly identify which publications would be included in this instance, to ensure the quality of the output.

However, the ERA Guidelines explicitly state that citation analysis assessments will only take into account indexed journal articles, so without a significant change being made to the fundamental metrics used for ERA, preprints will not be eligible for citation analysis anyway. From a practical perspective, it does not seem worth the substantial effort it would require to report all preprints when they are already precluded from being included in citation analysis.

3. Is it appropriate to include preprints as an eligible research output type for peer review disciplines? Yes/No. Please explain your answer.

No.

The same issues of ensuring the quality and uniqueness of the research outputs that are relevant in citation analysis assessments also apply to the peer review analysis assessment.

4. What would be the material advantages and/or disadvantages of including preprints in ERA?

Given that preprints are an accepted form of publication in some fields, inclusion of preprints in the data collection process would ensure the full gamut of research outputs are included in the ERA collection.

However, concerns remain about how to ensure the quality of preprint publications, as they have not undergone the full formal peer review process. To maintain the focus on research excellence, there would need to be very specific criteria to clearly identify which publications would be eligible, ensuring quality (a minimum threshold of peer review) and uniqueness (a DOI or specific URL assigned to the publication). There would also be challenges with ensuring consistency of data collection across university submissions.

5. What would be the impact upon universities if required to include all preprints in ERA and undertake deduplication to ensure research outputs submitted to ERA are unique?

Including preprints in the ERA collection would impose a significant additional administrative burden upon university resources. There is no standard mechanism to track or import these for analysis and evaluation.

Professor Mark Hutchinson
President, Science &Technology Australia

Misha Schubert
CEO, Science & Technology Australia