



Professor Emma Johnston AO FAA FTSE FRSN
Deputy Vice-Chancellor (Research)

9 November 2023

Rosie Hicks
CEO, Australian Research Data Commons

Via email: contact@ardc.edu.au

Dear Ms Hicks,

Draft Australian National Persistent Identifier (PID) Strategy

Thank you for the opportunity to comment on the first version of the ARDC National Persistent Identifier Strategy. Through the University of Sydney's close engagement with the sector and with the ARDC, we recognise the significant opportunities that are presented to the research sector by adopting a national strategy of harmonisation and standardisation of persistent identifiers (PIDs).

We are also very supportive of the consultative approach to the development of the strategy and pleased to be able to contribute our views to this process.

Summary

While we support the adoption of persistent identifiers for researchers, projects, data and instruments and recognise the value of doing so, the proposal in this strategy of limiting their use to research that is made FAIR only unnecessarily restricts their utility. Universal use of PIDs for all research purposes will enable researchers and institutions to, for example, identify and manage costs in research projects, track research activities, record all types of contributions to a project (including collaborations) and build context around research outputs.

The University endorses the designing of processes and systems that benefit **all** research.

Comments

We provide the comments below on the two main aims of the strategy:

1. Improve research quality and efficiency

<i>1. Increase the Findability, Accessibility, Interoperability and Reuse of inputs to research</i>
<i>2. Increase the Findability, Accessibility, Interoperability and Reuse of research outputs</i>
<i>3. Improve research reproducibility, provenance and attribution whilst minimising administrative burden, enabling researchers to spend more time on research</i>

While we agree with the vision statement and opportunities presented by the development of a national strategy, we are concerned that the strategy itself overly focuses on research that leads to data discovery and sharing (i.e. it is FAIR: Findable, Accessible, Interoperable, Reuseable).

1.1. Data discovery and sharing

We recognise the importance of increasing visibility and availability of research outputs for discovery and secondary use where it is appropriate to share this information, however, there are many confidential or otherwise sensitive outputs from research projects that cannot be shared at times, if at all.

For example, there will be publicly funded research projects containing highly sensitive data that will not be able to be shared and may even have restrictions on discovery, or require long embargo periods.

Increasingly, many of our researchers have commercial contracts with funding partners from industry, meaning that confidentiality and protection of Intellectual Property necessitate an embargo deferring or even prohibiting the wider sharing of research findings. Similarly, projects with international collaborations may have discovery and access restrictions beyond an initial short embargo period. Regardless, our university still needs to be able to manage, track and cost research beyond the proportion of projects that are FAIR.

We believe that the use of PIDs that will 'link researchers to projects, to grants, to organisations, to equipment' (part 3 of the strategy) is a necessity that needs to be applied to **all** research projects. We were disappointed to discover recently that RAiDs will be essentially restricted to FAIR research use only, as the introduction of a maximum embargo period (18 months) prevents their use for non-FAIR research projects. We had previously envisaged a use for RAiDs throughout our research administration systems to link and record PIDs covering all aspects of research activity, including use of both in-house and federated infrastructure in imaging, microscopy etc. We can, of course, implement our own unique project identifier for our internal systems, however, the absence of a nationally unique identifier for a project severely limits functionality in linking with NCRIS and other federated services e.g. microscopy, imaging, NCI etc, and recording contributions from inter-institutional collaborations, partnerships and grants.

We contend that restricting PIDs to 'FAIR data only' limits the efficiency and quality gains that 'improve research reproducibility, provenance and attribution' (part 3 of the strategy) that should be applied to all research at Australian universities and research institutions. We urge consideration of a broader and more inclusive approach to include research projects and outputs that cannot or should not be made FAIR.

1.2. Licensing

Given that this strategy is for improving discoverability and access to research outputs by government, business, industry and universities, we should be aware that this strategy may increase the use of university research data for commercial purposes. It is vital that appropriate licensing is used for published outputs, and that adequate resources are developed and provided to assist research institutions in applying these to research outputs.

1.3. Operationalisation

We stress the importance of the need to operationalise and manage PIDs, noting the infrastructure investment required to do so. The ARDC *modus operandi* has been to implement central systems to which IDs could be published and ignore the operationalisation of the entire process as something left to the universities to individually solve. If a national PID strategy is to have any impact, it needs to address the whole management cycle for PIDs, and not limit functionality to a webpage used to search for published PIDs only. In the case of Instrument IDs for example, creating PIDs is trivial; managing them is not, which is why previous efforts to track PIDs over the past decade have not succeeded.

It is also not clear how this strategy will increase the quality of research or research data. For example, linking data-creating instruments' IDs with datasets and projects contributes to provenance, however, unless it continuously integrates operational details of those instruments (calibrations, QC pipelines, upgrades etc) it has negligible benefit for research quality and impact, or wider societal impact. The link between this strategy and the ARDC's needs and objectives (e.g. relating to FAIR data) should be made clear.

2. Optimise the national research and innovation ecosystem

<i>4. Improve our ability to evaluate research quality, impact and evidence of public benefit</i>
<i>5. Optimise our ability to understand the impact of research inputs such as grants and investment in research infrastructure</i>
<i>6. Improve our ability to map Australia's research capability</i>

2.1. National metrics for research outputs

The assumption that all publicly funded research can and should be made FAIR in a short time period (embargo of 18 months for RAIIDs) undermines the objectives of parts 4, 5 and 6. While research outputs should be made FAIR where it is appropriate to do so, sensitive research projects cannot be shared and thus would not be provided with a PID. Consequently, sensitive research (without PIDs) will be omitted from metrics for quality, costings (including return on investment), and capability mapping.

We believe that universal application of PIDs for all research projects will maximise potential improvement and optimisation of processes and metrics at individual institutions. Metrics will be able to be collected for all research and could be contributed to national reporting capability.

2.2. Implementation of a national roadmap

Should the University actively participate in the development and implementation of a future PID roadmap, we will require extensive support for building PIDs into our infrastructure, systems, workflows, policies and training. We will need to have further discussion with ARDC, when appropriate, to understand how this will be resourced.

The development of a new strategy for PIDs presents the research sector with a rare opportunity to implement significant improvements in research quality and efficiency. We hope that the views expressed here can be incorporated into a national approach that will result in a sector-wide uplift in the management and reach of our research.

Thank you again for this opportunity to provide feedback on the draft strategy and if further explanation of any of our comments is required, please do not hesitate to contact Adele Haythornthwaite via adele.haythornthwaite@sydney.edu.au.

Yours sincerely,

(signature removed)

Professor Emma Johnston
Deputy Vice-Chancellor (Research)