



**Australian Government**  
**Australian Research Council**

**AUSTRALIAN RESEARCH COUNCIL SUBMISSION (ARC)**

**to the draft report**

**PRODUCTIVITY COMMISSION INQUIRY INTO DATA AVAILABILITY AND USE**

**Introduction**

This second submission follows that made by the ARC in July 2016 and is in response to the Productivity Commission Draft Report Data Availability (the Draft Report). The submission contains further information and should be read in the context of the first submission.

In particular, this submission provides further information in respect of draft recommendations 3.1, 3.2, 9.6, 9.7, 9.8 and 9.9 as well as Section 3 (*Public sector and research data collection and access*), Section 5 (*It's all about you: the challenges of using identifiable information*) and Section 10 (*Trust: the foundation of the new data framework*).

**General Comments**

The ARC notes the potential of increased data availability and use identified by the Commission in its draft report, and it is committed to contributing to maximising these benefits.

As indicated in its submission the ARC is committed to improving the availability and use of the public sector data that it holds. The ARC notes that greater access to publicly held data needs to be commensurate with the cost of that access. The ARC is of the view that the report does not adequately recognise the cost that goes with open access to data.

The resourcing issues regarding many of the recommendations contained in the draft report are not adequately addressed. The creation of comprehensive, easy to access data registers is not a simple exercise, and without adequate resourcing would not be achievable by 1 October 2017.

The ARC is supportive of making data from publicly funded research bodies publicly available, however it questions the need for the proposed National Data Custodian to grant, on an approved project-specific basis, trusted user access to personnel from a range of potential entities, including: all Australian Government and state and territory government agencies; all Australian universities; and other entities (be they corporations, not-for-profit organisations or research bodies). The approach outlined on page 22 of the draft report regarding providing greater access to data of publicly funded research beyond the initial researchers, including key features such as granting access on a project-specific basis to approved personnel in either Australian or state/territory government agencies and to approved researchers appears overly complex. Making open access the default position for publicly-funded research data outputs would seem to remove any need for approved access and use by the National Data Custodian (NDC) and Accredited Release Authorities (ARAs).

The ARC is referenced a number of times throughout the draft report, often as the sole example of identified issues around the availability of research-related data. The report does not reflect that the ARC provides less than 10 per cent of the Australian Government's investment in research and

development (R&D). The Productivity Commission may wish to draw on other funders of research as examples in its report.

The ARC commitment to maximising the benefits from research data arising from ARC funded research, as well as other data sets it holds, is not accurately represented. Since 2007, the ARC has encouraged researchers to deposit data arising from ARC funded research projects in publicly accessible repositories. Since February 2014, the ARC has required researchers to outline how they plan to manage data arising from their research.

The ARC already makes a significant amount of data on its funded research available and is working on making a larger amount of its Excellence in Research for Australia (ERA) data available.

Whilst there is a significant section of the report dealing with issues around health data, as well as considering that the NHMRC provides more public funding for research than the ARC, mentions of the NHMRC, both in number and context, do not feature as significantly as the ARC.

In simple terms the ARC would be able to enact a policy that mandates making ARC-funded research data publicly available. However, such a requirement would not necessarily adequately reflect the particular approaches, standards and uses for data that may exist in different institutions, disciplines and research projects. Whilst some institutions and disciplines may have infrastructure and/or processes in place for storing, managing and sharing data, which are valuable resources that should be utilised, others may not have such resources available and would need to be resourced to do so.

## **SUGGESTED AMENDMENTS and COMMENTS:**

There are a number of instances where references to the ARC in the draft report require adjustment.

1. The ARC linkage project listed on page 6 and 194 (Section 5 – Box 5.1) should be edited:
  - ARC funding for the project (LP0882394) associated with this study ceased in 2010. Suggested wording change to references on page 6 and 194:  
  
*In 2008, the Australian Research Council<sup>1</sup> and other government bodies provided funding for the Vaccine Assessment Using Linked Data Safety Study.....*
2. The ARC is the only entity listed under the title of Lack of discoverability of research page 140 and in draft recommendation 3.2. The text should be amended to reflect that there are other significant funders of research in Australia such as the NHMRC.
3. Page 136 of the draft report raises that attempts to monitor data availability are impeded by the lack of a registry of datasets that result from Australian Research Council (ARC) funding.
  - The ARC wishes to stress that it is not resourced to develop and maintain such a registry. The commentary does not provide an accurate representation of the devolved system of research management in Australia and the role of universities in managing and making available their data outputs.
  - The commentary is also not representative of the ARC's commitment to maximising the benefits from research data arising from ARC funded research, as well as other data sets it holds. Since 2007, the ARC has encouraged researchers to deposit data arising from ARC funded research projects in publicly accessible repositories. Since

---

<sup>1</sup> ARC funding for this project was from 2008 - 2010

February 2014, the ARC has required researchers to outline how they plan to manage data arising from their research.

4. To evidence the lack of a registry of datasets that result from Australian Research Council (ARC) funding the draft report lists a single metric from the ERA 2015 National Report. The referencing of this metric is somewhat misleading.
  - Figure 3.5 is taken from the ERA 2015 Report – National Overview page 86.
  - The figure is in reference to open access to publications, not open access to data.
  - There is no detail provided as to why open access (OA) is low as is explained in the ERA report (i.e. the ARC's OA policy only took effect from 1 Jan 2013 which is after the reference period for the ERA 2015 Report; the institutions in many cases answering the OA status of outputs retrospectively; and that the data should be treated with caution).
  - The ARC did not provide any information in its submission regarding this study.
  - ARC requests that the following caveat be added should the Commission wish to include figure 3.5 in the final report:

*The information gathered regarding open access to research publications in the ERA 2015 submission process, required Eligible institutions to answer yes or no as to whether a research output had been made available in an open access repository. Therefore, the information should be treated with caution and will only form a baseline for more accurate reporting in any future rounds.*

5. The ARC is listed as a sole example in bullet point two within the Academic institutions heading on page 379. The ARC is not an academic institution. It is a funder of research. As mentioned previously the ARC provides less than 10 per cent of the Commonwealth Government's investment in research and development (R&D) and the report would benefit from recognising that there are other funders of research at academic institutions.
  - The ARC does not commission datasets or generate research data sets.
  - Since 2007, the ARC has encouraged researchers to deposit data arising from ARC funded research projects in publicly accessible repositories.
  - Since February 2014, the ARC has required researchers to outline how they plan to manage data arising from their research.
  - The ARC already publishes up-to-date datasets of the research that it funds and much of the data collected as part of its Excellence in Research for Australia (ERA) initiative.

## **ARC COMMENTS ON SPECIFIC RECOMMENDATIONS**

### **Draft Recommendation 3.1**

ARC comment:

- The Draft Report proposes a framework for Australia's data future which is not fully reflective of the many types of data produced and the different standards and approaches to data management, analysis, archiving, storing, sharing and publication.

- Whilst the ARC is in support of the draft recommendation, the due date of 1 October 2017 provides only a short window for implementation and without any reference to the resourcing that would be required.
- This compressed timeframe, without resourcing, is likely to create significant risks to the end users if datasets are released before proper consultation and framework is in place, relevant rules on extraction are provided and the format of the data is in place.

### **Draft Recommendation 3.2**

ARC comment:

- The ARC is not a publicly funded research entity. It is a public funder of research.
- As stated in our submission, the ARC provides approximately \$800 million in research funding through the National Competitive Grants Programme each year. While significant, this represents less than ten per cent of Australian Government investment in research and development.
- The ARC suggests editing the recommendation as follows:

*Publicly funded research entities, such as CSIRO, as well as funders of public research, such as the National Health and Medical Research Council and the Australian Research Council, should publish up-to-date registers of data holdings, including metadata, that they fund or hold.*

### **Draft Recommendations 9.6 - 9.8**

- The ARC is supportive of making data from publicly funded research bodies publicly available, however it questions the need for the proposed National Data Custodian to grant, on an approved project-specific basis, trusted user access to personnel from a range of potential entities, including: all Australian Government and state and territory government agencies; all Australian universities; and other entities (be they corporations, not-for-profit organisations or research bodies).
- Whilst access would be granted for the life of the specific approved project etc. would it not be simpler to make open access the default position for publicly-funded research data outputs (refer to recommendation 9.9) and remove any need for approved access and use by the National Data Custodian (NDC) and Accredited Release Authorities (ARAs).

### **Draft Recommendation 9.9**

- In simple terms the ARC would be able to enact a policy that mandates making ARC-funded research data publicly available. However, such a requirement would not necessarily reflect the particular approaches, standards and uses for data that may exist in different institutions, disciplines and research projects.
- While some institutions and disciplines may have infrastructure and/or processes in place for storing, managing and sharing data, which are valuable resources that should be utilised, others may not have such resources available and would need to be resourced to do so.