

Australian National Data Service submission to the Productivity Commission Draft Report on Intellectual Property Arrangements.

Productivity Commission DRAFT RECOMMENDATION 15.1

All Australian, and State and Territory Governments should implement an open access policy for publicly-funded research. The policy should provide free access through an open access repository for all publications funded by governments, directly or through university funding, within 12 months of publication. The policy should minimise exemptions.

This submission by the Australian National Data Service¹ (ANDS) argues for expanding the scope of R15.1 beyond just publications, to include the other major outputs of research: research data, techniques, algorithms and software. Such an expansion would make R15.1 consistent with current Australian Government policy, policies of the NHMRC and ARC as well as those of most overseas funders. Collectively, research data, techniques, algorithms and software (sometimes referred to as non-traditional research outputs or NTRO) are increasingly being considered as key enablers of innovation and especially in relation to improving the collaboration between researchers and business and industry.

An important note: publications, as opposed to data etc (NTRO) have been the subject of open access policies by Australia's major funders for a number of years now. This submission seeks to extend this beyond publications to also include data, techniques, algorithms and software, thus bringing Australia into line with the major overseas funders.

1. Background

ANDS is transforming Australia's research data environment to make Australian research data collections more valuable by managing, connecting, enabling discovery and supporting the reuse of this data. Data reuse drives innovative research and supports policy development; it also benefits the wider community, including business and industry. Effective and efficient data reuse can only happen with the right policy settings. The economics of data sharing and reuse are overwhelmingly attractive; a recent Australian study (now replicated several times overseas) by Houghton and Gruen², using conservative techniques, estimate that the value of data in Australia's public research to be at least \$1.9 billion and possibly up to \$6 billion a year at current levels of expenditure and activity.

Draft recommendation 15.1 (*open access policy for publicly-funded research*) seems to be focused on only one component of research outputs, viz. Publications. Should R15.1 be expanded to include all of the major outputs (i.e. publications plus NTRO), the benefits to

¹ <http://www.ands.org.au/>

² http://www.ands.org.au/__data/assets/pdf_file/0019/393022/open-research-data-report.pdf

Australia could be very significant indeed and would be entirely consistent with the Australian Government's National Science and Innovation Agenda³.

2. Most international open access policies include NTR0

In February 2013, the Obama Administration took an important step toward increasing public access to the results of research funded by the Federal Government, focusing on two key products of funded research: peer-reviewed scholarly publications and scientific data. The rationale for these policies is the recognition of the potential wider benefit of research data: "Policies that mobilize these publications and data for re-use ... will accelerate scientific breakthroughs and innovation, promote entrepreneurship, and enhance economic growth and job creation."⁴

In June, 2013, the G8 Science Ministers issued a statement endorsing the value of open data as supporting innovation and helping to solve global challenges: "We are committed to openness in scientific research data to speed up the progress of scientific discovery, create innovation, ensure that the results of scientific research are as widely available as practical, enable transparency in science and engage the public in the scientific process."⁵

There is also a trend among research funders allowing research data to be considered a normal part of research outputs⁶ which is based on an understanding of the inherent value of research data in the scientific method as well as its potential applied value in supporting collaboration, innovation, public policy, industry, education, public engagement, and in helping to address larger national and global challenges.

In the USA, both the NSF and NIH require data sharing plans as part of the application process. The NIH position is summarized below:

"Data sharing is essential for expedited translation of research results into knowledge, products and procedures to improve human health. The Final NIH Statement on Sharing Research Data was published in the NIH Guide on February 26, 2003. This is an extension of NIH policy on sharing research resources, and reaffirms NIH support for the concept of data sharing. The new policy becomes effective with the October 1, 2003 receipt date for applications or proposals to NIH".⁷

Similar policies exist in the UK, the EU and in many other countries.

³ <http://www.innovation.gov.au/>

⁴ Increasing Access to the Results of Federally Funded Scientific Research :

http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

⁵ G8 Science Ministers Statement <https://www.gov.uk/government/news/g8-science-ministers-statement>

⁶ This is expressed through policies like those of the NSF which allow citable data products as one of the elements of a researcher's track record when applying for competitive grants. See: http://www.nsf.gov/pubs/policydocs/pappguide/nsf13001/gpg_2.isp#IIC2fic

⁷ http://grants.nih.gov/grants/policy/data_sharing/

In the last few years the issue of scientific reproducibility is growing in importance, such that access to the research data is necessary and in many cases, mandatory.

3. Journals increasingly mandating access to research data

In recent years, journals are increasingly requiring 'click-through' access to the data sets which support the publications. In some cases (after some high-profile retractions) the provision of data is mandatory. Examples include: Nature⁸, Science⁹ and PLOS Medicine¹⁰. It seems likely that this trend will continue to spread to the publication sector more generally. The following link from Victoria University lists many other journals requiring data deposit: <http://guides.library.vu.edu.au/content.php?pid=489543&sid=4015042>

4. Australian Government Policies

In December 2015 the Australian Government released two related initiatives, the National Science and Innovation Agenda (NISA) and the Public Data Policy Statement¹¹ (PDPS). Prior to this, the Australian Government's position on data (meaning both government data and research data) was fragmented across agencies. The underlying premise of the new policies is that data which has been paid for using public money is now to be considered an asset with potential benefits for researchers, business and beyond. The PDPS also recognises those benefits cannot be fully realised without proper data management, standards, licences, repositories and services to ensure the data can be discovered, shared and reused effectively.

NISA has many references to data and the opportunities around its clever reuse. The PDPS recognises the potential for innovation which can only be realised by increasing access to public data, including both the data behind the administrative functions of government as well as the data that comes from publicly funded research. This quote from the PDPS outlines the importance of the other (non-publication) outputs: "Australia's capacity to remain competitive in the digital economy is contingent upon its ability to harness the value of data."

5. Australian Funders' policies

Australian policies also highlight the importance of research data as an output of research. The Australian Code for the Responsible Conduct of Research says: "The potential value of the material for further research should also be considered, particularly where the research would be difficult or impossible to repeat."¹² The ARC policy states "The Final Report must outline how data arising from the Project has been made publicly accessible where appropriate" (Discovery Projects Funding Rules for funding commencing in 2014). These policy settings are entirely consistent with NISA which emphasizes the increasing

⁸ <http://www.nature.com/authors/policies/availability.html#data>

⁹ http://www.sciencemag.org/site/feature/contribinfo/prep/gen_info.xhtml#dataavail

¹⁰ <http://journals.plos.org/plosmedicine/s/data-availability>

¹¹ https://www.dpnc.gov.au/sites/default/files/publications/aust_govt_public_data_policy_statement_1.pdf

¹² Australian Code for the Responsible Conduct of Research, Section 2: Management of Research Data and Primary Materials, Introduction. <http://www.nhmrc.gov.au/files/nhmrc/publications/attachments/r39.pdf>

importance of data outputs in research, "A unifying thread across all the domains is the importance of data. Australia's key research challenges will be increasingly data intensive and data driven."¹³

The Australian Research Council has recently announced changes to the funding rules for schemes under the Discovery Program for 2014 and 2015 and similar changes are expected for relevant Linkage Schemes under the National Competitive Grant Program (www.arc.gov.au/ncgp/). These changes will have implications for how research data is managed. The new rules strongly encourage researchers to make their data available: "Researchers and institutions have an obligation to care for and maintain research data in accordance with the Australian Code for the Responsible Conduct of Research (2007). The ARC considers data management planning an important part of the responsible conduct of research and strongly encourages the depositing of data arising from a Project in an appropriate publicly accessible subject and/or institutional repository".¹⁴

The NHMRC's 2015 data sharing statement¹⁵ includes this quote: "NHMRC encourages data sharing and providing access to data and other research outputs (metadata, analysis code, study protocols, study materials and other collected data) arising from NHMRC supported research. This aligns with researchers' responsibilities under the Australian Code for the Responsible Conduct of Research (2007), which provides advice on the storage, management and privacy of research data (section 2.5-2.7) and states: "Research data should be made available for use by the other researchers unless this is prevented by ethical, privacy or confidentiality matters."

6. Recommendation 15.1 (proposed extended version)

ANDS hopes that it has provided the Productivity Commission with sufficient evidence for it to consider that R15.1 should be re-scoped or extended beyond publications to include publications (the 'traditional' form of research outputs) plus research data, techniques, algorithms and software (the 'non-traditional' outputs). There are compelling arguments for doing so which related to fostering research and innovation, national and international policy and norms, as well as overwhelming economic arguments. Increasingly, the global issue of research reproducibility¹⁶ also demands access to research data underpinning publications.

A suggested rewording of R15.1 follows:

All Australian, and State and Territory Governments should implement an open access policy for publicly-funded **research outputs, including publications, research data, techniques, algorithms and software**. The policy should provide free access through an open access repository for all publications funded by

¹³ 2012 National Research Investment Plan p 61 <http://www.innovation.gov.au/research/Documents/NationalResearchInvestmentPlan.pdf>

¹⁴ <http://www.arc.gov.au/pdf/DP15/Funding%20Rules%20for%20the%20Discovery%20Program.pdf> (p.18)

¹⁵ <https://www.nhmrc.gov.au/grants-funding/policy/nhmrc-statement-data-sharing>

¹⁶ <http://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>

governments, directly or through university funding, within 12 months of publication. The policy should minimise exemptions.

Concluding comments: open access to publications has been the norm in Australia for a number of years. In principle, the 'same' open access policies can apply to data, but at the present time this is very limited. So it is not the open access policies *per se* that need amending, it is the scope of what is considered research outputs. R15.1 (amended) would bring Australia into line with major international policy on publicly-funded research.

For further information please contact Adrian Burton