

The Australian Academy of Science welcomes the National Archives of Australia's (NAA) consultation on the exposure draft of the policy *Building trust in the public record: managing information and data for government and community* (the Policy). This submission draws on the expertise of the Academy's National Committee for Data in Science; Professor Michael Barber FAA FTSE, Chair, Australia Research Council Learned Academies Special Project on 'Big Data in Australia Research: Issues, challenges and opportunities'; and Dr Danny Kingsley, Scholarly Communication Consultant.

The Academy has a strong interest in information management policy, particularly in the context of research data, and supports the NAA's aim to enable continuous improvement in managing Australian Government information assets. It is overwhelmingly accepted that data is valuable to both the private and public sectors. The information assets held by the Australian Government enable public policy issues to be assessed and addressed from a range of perspectives. As such, the NAA and *Archives Act 1983* are of great interest to the Australian science sector.<sup>1</sup>

In this submission, the term 'research data' is used to capture both *data that is used for research*, which could be historical data, and *data that is generated from research*, which could become historical research data, since both forms apply to Government information assets. For example, historical climate data collected by the Bureau of Meteorology is frequently used by researchers outside of Government, and data generated by research programs funded by the Australian Government can be useful for informing Government research and policy activities.

### *FAIR and CARE Principles*

The Academy encourages the NAA to ensure that any changes in its policy uphold the principles of:

- FAIR (Findable, Accessible, Interoperable and Reusable): Supporting open data practices to facilitate increased data sharing;<sup>2</sup> and
- CARE (Collective Benefit, Authority to Control, Responsibility and Ethics): Supporting the right for Indigenous People to govern the collection, ownership and application of their data.<sup>3</sup>

The 'FAIR Guiding Principles' and 'CARE Principles for Indigenous Data Governance' are complementary and encourage custodians and users of research data to consider both people and purpose.

Research data is generated, used across and shared between both government and research sectors. The Academy recommends policies be developed that make it easier for data to be shared and to be interoperable across both sectors, i.e. information management policies for both sectors should be aligned.

### *Fit-for-purpose – Data is accessible and re-usable*

There is value in the research sector being able to access information assets held by the Australian Government. Importantly, given that digital technologies are constantly evolving, such information needs to be machine-readable. Machine-readable formats will assist information remaining fit-for-purpose and usable in the future. This is pertinent as future applications of information are not always predictable and are challenging to identify in advance. Where the Policy states that “sustainable digital file formats are used to ensure content is available for as long as needed” (page 10), it is important to note that, in some cases, this may be in perpetuity, and appropriate information management processes must be considered and implemented.

The Academy supports the Policy’s suggestion that stored “information and assets have adequate and standardised descriptive information (metadata)”. Obtaining accurate metadata is imperative to managing information according to the FAIR and CARE principles.

The importance of information management to enable its re-use, which includes the ease at which information can be found and understood, is reiterated throughout the Policy. This is also noted in NAA’s Information Management Standard for Australian Government, which states that information “can be found, retrieved, and interpreted when needed”.<sup>4</sup> On page 15, the Policy recommends that Government agencies “implement strategies, including storage and preservation strategies, for the management of all information assets”. The Academy recommends that this be mandatory: if preservation strategies are not in place from the beginning, information assets will quickly be lost through inaccessibility. Good digital preservation underpins information asset management.

However, in NAA’s 2019 survey that measured agencies’ progress in managing information assets, a capability gap “interoperability of data between systems” was identified (page 8). A suggestion to enhance this capability is to ensure that the Policy is explicit and emphasises that access to and re-use of stored information applies to both human-readable and machine-readable formats. Data must be curated according to the FAIR principles, including machine-actionable standards before it is archived. As noted above, this is crucial for information assets that are to be maintained for long periods of time, e.g. longitudinal studies.

Further to this, on page 8 of the Policy, a recommendation is made for Australian Government agencies to register their information assets and document their value. The Policy should be explicit in listing where information assets can be registered and what decision-making processes are in place to determine if an information asset has value.

Interestingly, the Australian Government’s data.gov.au website provides a ranking for datasets and takes into consideration:<sup>5</sup>

- Is the data available on the web, and does it have an open license?
- Is the data available as structured, machine-readable data?
- Is it available in a non-proprietary format?
- Does the data declare a schema?
- Does the data link to other data?

The Academy wishes to highlight that in addition to the above considerations, the NAA should also insist that data is available in its raw format or with minimal manipulation or conversion. This is particularly important for research applications.

### *Data is securely stored*

Many Australian Government agencies generate and hold onto vast amounts of data with significant potential to inform policy development and contribute to Australia's economic growth. Storage security of this data is imperative to achieving community trust.

The Academy would like to draw the NAA's attention to the Five Safes Framework, which is designed to reduce the risk of sensitive data being misused.<sup>6,7</sup> The principles of the framework are Safe People, Safe Projects, Safe Settings, Safe Data and Safe Outputs. The framework was developed by the United Kingdom's Data Service and provides guidance on how to design, implement and assess data systems to mitigate data mismanagement and consequent harm.

The Office of the National Data Commissioner (ONDC) have developed their Data Sharing Principles,<sup>8</sup> based on the UK's Five Safes Framework, which is designed to enable safe and appropriate data sharing:

1. Projects: Data is shared for an appropriate purpose that delivers public benefit
2. People: The user has the appropriate authority to access the data
3. Settings: The environment in which the data is shared minimises the risk of unauthorised use or disclosure
4. Data: Appropriate and proportionate protections are applied to the data
5. Output: The output from the data-sharing arrangement is appropriately safeguarded before any further sharing or release.

As noted in the Policy, achieving proper information management will maximise the value of the information assets. Poor information management can result in the loss of information assets, this risk is in addition to those noted on page 7. Regenerating lost information assets can be costly and, in some cases, may not be possible. Proper management processes will not only maximise the value of information assets but will also ensure that invaluable information assets are stored in a sustainable manner.

To discuss or clarify any aspect of this submission, please contact Mr Christopher Anderson, Director of Policy, Australian Academy of Science ([Chris.Anderson@science.org.au](mailto:Chris.Anderson@science.org.au)).

## References

1. Australian Government. Archives Act 1983. *Federal Register of Legislation*  
<https://www.legislation.gov.au/Details/C2019C00005>.
2. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Nat. Sci. Data* **3**, 1–9 (2016).
3. Global Indigenous Data Alliance. CARE Principles of Indigenous Data Governance.  
<https://www.gida-global.org/care>.
4. National Archives of Australia. Information management standard for Australian Government. <https://www.naa.gov.au/information-management/information-management-standards/information-management-standard-australian-government>.
5. Australian Government. Linked data rating - beta. *data.gov.au*  
<https://data.gov.au/page/linked-data-rating>.
6. Desai, T., Ritchie, F. & Welpton, R. Five Safes: Designing data access for research. (2016).
7. UK Data Service. Regulating access to data. <https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes>.
8. Commonwealth of Australia - Department of the Prime Minister and Cabinet. Best practice guide to applying data sharing principles. (2019).