

President Professor John Shine AC PresAA FRS



**Australian
Academy of
Science**

Inquiry Secretary
Senate Finance and Public Administration Committees
PO Box 6100
Parliament House
Canberra ACT 2600

By email: fpa.sen@aph.gov.au

Friday, 12 March 2021

Dear Secretary,

Australian Academy of Science submission on the Data Availability and Transparency Bill 2020

The Australian Academy of Science (the Academy) welcomes the opportunity to address the *Data Availability and Transparency Bill 2020*, currently subject to an inquiry by the Senate Finance and Public Administration Legislation Committee.

The Academy has a strong interest in information management policy, particularly in the context of data for research and subsequently, how this research can be used to inform and support public policy. This is acknowledged in the Bill as one of its three purposes for data sharing, which are: (i) delivering government services, (ii) design and implementation of government services and (iii) research and development. This submission focuses on the benefits from establishing a framework for the sharing of data for the purposes of scientific inquiry, review and to advance knowledge.

The Academy welcomes the establishment of a National Data Commissioner (the Commissioner) and a National Data Advisory Council, which will include the Chief Scientist of Australia as an *ex officio* member, who will oversee the accreditation process that will provide a robust and safe mechanism for the sharing of Commonwealth government data. The Bill before the Senate enables government data to be shared with accredited users for specific purposes that are in the public interest. If implemented appropriately, the Bill should enable researchers to continue to use data in ways that benefit Australians, safely and responsibly.

In the context of data access for research, the Academy would like to draw the Committee's attention to the below points:

Access to data

One of the primary purposes of the legislation is to support research and development. The Academy strongly recommends that, for research purposes, the Bill must specifically refer to the FAIR (Findable, Accessible, Interoperable and Reusable) Principles for research data management.¹ The FAIR Principles allow for regulated access and, if implemented correctly, require data management plans to articulate a clear pathway to access and that access is not unnecessarily or unreasonably withheld.

Access to data is essential for its value to be realised. The research community has requirements for FAIR data management and Government departments collecting data should follow the same data management principles. The importance for FAIR data management in research is reflected in the 2018 National Collaborative Research Infrastructure Strategy (NCRIS) guidelines and the Australian Research

Data Commons' guidelines for project data outputs.^{2,3} In line with this, Clause 77, Criteria for accreditation, should be explicit in requiring the custodians of Government data to manage public sector data according to the FAIR Principles.

Accessible data also means collecting, storing and sharing data in machine-readable formats, which will ensure that data assets remain fit-for-purpose and useable in the long-term. The Explanatory Memorandum states that Clause 77 will ensure that "data can meet evolving technology, privacy and security requirements... into the future". This is again reiterated as the Bill enables the Commissioner to update approved forms to cater for future needs (Clause 132).

We note that, although the CARE (Collective Benefit, Authority to control, Responsibility and Ethics) Principles for Indigenous Data Governance are *not* explicitly stated, the Explanatory Memorandum does acknowledge that the Commissioner will "address cultural barriers to [the] sharing" of data.⁴

However, the Academy has identified an area for potential concern regarding 'data minimisation' (Clause 16 – Data sharing principles, paragraph 131). We realise that in some cases, the full set of data cannot be provided and a subset or a treated subset is provided instead. However, metadata explaining how the subset was derived and what treatment was applied to the data are required, and missing subsets should be kept on file. Metadata provides context for how data, or a subset thereof, was generated, updated or maintained. In the context of research, this information is vital to supporting researchers' efforts to deliver reproducible research outputs.

Regulated access to the full set of unprocessed data is often necessary, particularly access for the purpose of reviewing research data and maintaining research integrity thereby building confidence to use research findings in public policy. A recent example testifies to the importance of reviewers having access to the full set of data: a potential COVID-19 therapeutic was reported to increase the risk of death in hospitalised patients and international trials were halted. It was only when reviewers requested access to the research data, and their request was declined, that the data was found to be inadequate and the publication was subsequently retracted and clinical trials resumed.^{5,6}

Social licence to operate

The Explanatory Memorandum notes that the Bill was developed using a 'privacy by design approach', which is in line with the Five Safes Framework for data and risk management. This approach is vital to ensuring that the Bill supports the Commissioner in gaining public trust and developing the social licence to securely share Government data for public good.

Compliance by accredited entities will ensure that the integrity of the scheme is maintained. The Bill will support the Commissioner in having regulatory power to monitor and enforce the requirements of the scheme. Having monitoring and enforcement processes in place will be crucial for maintaining public trust.

The Bill acknowledges existing mechanisms and resources may be applicable regarding ethics approval processes for the scheme. This acknowledgement will prevent duplication and unnecessary administrative burden, particularly where research projects have already undergone ethics reviews. However, it is worthwhile recognising that not all ethics committees are adequately equipped to manage data-specific issues: studies that involve access to complex data warrant expert analysis at the access approval stage. For example, the recent case outlined above could have been avoided if a more expert

review of the data had preceded the study. Data-specific expertise is also required when a dataset is to be used to train artificial intelligence or machine learning algorithms; the data must be assessed for its appropriateness e.g. that it is representative of a population and any underlying biases noted.^{7,8} Thus, when acknowledging existing processes, the approval process must determine if, in existing processes, the relevant expertise was available to assess data-specific issues.

The Academy believes the Bill is supportive of research and development activities, the outputs of which can also be used to inform Government policy and programs and advance the wellbeing of Australians.

To discuss or clarify any aspect of this submission, please contact Mr Chris Anderson, Director Science Policy at .

Yours sincerely,

Professor John Shine AC PresAA FRS

President

The Australian Academy of Science

References

1. FAIR principles. *GO FAIR* <https://www.go-fair.org/fair-principles/>.
2. Training, A. G. D. of E. and. Page 23: National Collaborative Research Infrastructure Strategy (NCRIS) 2018.
3. Australian Research Data Commons. FAIR data guidelines for project data outputs. https://ardc.edu.au/about_us/policies-and-guidelines/fair-data-guidelines-for-project-data-outputs/.
4. CARE principles of Indigenous data governance. *Global Indigenous Data Alliance* <https://www.gida-global.org/care>.
5. Mehra, M. R., Desai, S. S., Ruschitzka, F. & Patel, A. N. RETRACTED:Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet* **0**, (2020).
6. Mehra, M. R., Ruschitzka, F. & Patel, A. N. Retraction—Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *The Lancet* vol. 395 1820 (2020).
7. Challen, R. *et al.* Artificial intelligence, bias and clinical safety. *BMJ Qual. Saf.* **28**, 231–237 (2019).
8. The Australian Academy of Health and Medical Sciences. *Artificial intelligence in health: exploring the opportunities and challenges*. (2020).