AUSTRALIAN GOVERNMENT DATA SHARING AND RELEASE LEGISLATION

ISSUES PAPER

SUBMISSION BY

Associate Professor
Leanne Wiseman
Griffith Law School
Nathan campus
Brisbane

Associate Professor
Jay Sanderson
USC
Sunshine Coast

We thank you for the opportunity to comment on the Issues Paper for the proposed **Australian Government Data Sharing and Release Legislation.**

Given the short time frame for comments, rather than addressing the specific questions outlined in the Issues Paper, we have addressed our comments at the broader aims, goals and potential challenges that may arise from the implementation of the proposed data governance reforms. We would welcome the opportunity to contribute further to the discussion of the data governance reforms as they are developed.

It is important to place this current discussion in the context of the open data movement more generally.

"Open Access by default"?

As was noted in the Issues Paper (p. 3):

> In late 2015, the Prime Minister released the Australian Government Public Data Policy Statement. This statement provides a mandate for Australian Government entities to optimise the use and reuse of public data to drive innovation across the economy. This includes mandating the release of data as open by default when it is safe to do so.

This move to open access to Government and publicly funded data was part of a world-wide movement that followed closely onto the opening of publicly funded research. Large philanthropic funders of research, such as Gates Foundation, UK Research Council, CGIAR, World Bank etc all adopted similar policies from 2013-2015.

What is interesting to note is that in the past 5 years, the legal and practical difficulties experienced by many publicly funded research organisations to operationalise their open access mandates, has led many to rethink the 'open by default' approach to ensuring that once the data is made openly accessible, that it is also interoperable and reusable and so that the potential benefits of open data can be realised. For example, the CGIAR (formerly known as the

Consortium of International Agricultural Research Centres) one of the largest funders of public good research, has adopted an Open Access and Data Management policy predicated on the idea that data is "open as possible, closed as necessary"; this clearly acknowledges that it is not always best to have simply open data. This clearly acknowledges the legal restrictions of IP rights and confidentiality and contractual restrictions and that much of the data that has been collected and aggregated historically has not been done with the full and open consent of data contributors to on-sharing.

The F.A.I.R Principles

While it is noted on page 7 of the Issues Paper that the:

> DS&R Bill will apply appropriate and consistent safeguards to data sharing and release... [t]hrough the internationally recognised *Five-Safes disclosure risk management framework*, where safeguards can be dialed up or down as appropriate depending on the sensitivity of the data and whether privacy needs to be maintained…

we consider that the additional principles that the data be "findable", "accessible", "interoperable" and "usable" (F.A.I.R) also be part of the overall consideration of how data should be shared and released. Sharing and releasing data will not make the data more useable; the data that is released must be firstly legally able to be released (i.e. free from contractual and IP restrictions, properly curated, interoperable).

As the Australian National Data Service (ANDS) notes the F.A.I.R principles, published in 2016, have since received worldwide recognition by various organisations including FORCE11, National Institutes of Health (NIH) and the European Commission who now use this as a useful framework for thinking about sharing data in a way that will enable maximum use and reuse.[1] To illustrate the approach being taken by many international public funding organisations, further details of the F.A.I.R. principles are provided below.

> Based on these 15 principles, a set of **14 metrics** have been defined to quantify levels of FAIRness. The latest developments on FAIR are available at **GO-FAIR**.
>
> ***TO BE FINDABLE:***
>
> F1. (meta)data are assigned a globally unique and eternally persistent identifier.
> F2. data are described with rich metadata.
> F3. (meta)data are registered or indexed in a searchable resource.
> F4. metadata specify the data identifier.
>
> ***TO BE ACCESSIBLE:***
>
> A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
> A1.1 the protocol is open, free, and universally implementable.
> A1.2 the protocol allows for an authentication and authorization procedure,

[1] https://www.ands.org.au/working-with-data/fairdata

where necessary.
A2 metadata are accessible, even when the data are no longer available.

TO BE INTEROPERABLE:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles.
I3. (meta)data include qualified references to other (meta)data.

TO BE RE-USABLE:

R1. meta(data) have a plurality of accurate and relevant attributes.
R1.1. (meta)data are released with a clear and accessible data usage license.
R1.2. (meta)data are associated with their provenance.
R1.3. (meta)data meet domain-relevant community standards.[2]

These principles ensure that data that is shared and released, or made open, is in fact in a format that is usable. There is much work being done internationally to develop ontologies and semantics to assist in the process of making data that is released interoperable and ultimately reusable (i.e. F.A.I.R.).[3]

The overlap with privacy law

As the Productivity Commission noted, there is a commonly held belief that "….information about an individual will be personal information and thus protected under the Privacy legislation in Australia." They also, however, noted the complexities of, and gaps in, the legal and regulatory frameworks around data access and use, noting that, in Australia:

> Legislation restricting access to data was formulated up to a century ago, and much is no longer fit for purpose. The primary legal impediment to more effective use of data is typically not the Privacy Act, but regulations and guidelines specific to the field in which the data is collected.[4]

and:

> A wide range of more than 500 secrecy provisions in Commonwealth legislation plus other policies and guidelines impose considerable limitations on the availability and use of identifiable data. While some may remain valid, they are rarely reviewed or modified. Many would no longer be fit for purpose.[5]

---

2 https://www.force11.org/group/fairgroup/fairprinciples

3 See Research Data Alliance https://www.rd-alliance.org

4 Australian Productivity Commission, 'Data Availability and Use' (Inquiry Report No 82, 31 March 2017) 121.

5 Ibid 133.

Therefore, it is essential that the DR&R scheme helps clarify this complexity, not merely add to it. Specifically, thorough consideration needs to be given to the way the privacy scheme in Australia would interact with the DR&R legislation. For example: Australia's current privacy legislation provides protection to individuals against misuse of their "personal information' by Government and Government entities. How does "data" or "personal data" relate to "personal information"?

Further, what has been an interesting trend in the past 5-10 years has been the change in language from "personal information" to "personal data". Indeed, in many contexts, the language of "personal information" seems now to be overtaken by the language of personal data. The recent implementation of the *General Data Protection Regulation* (GDPR) in the EU, highlights that the protection, privacy and security of personal data has now become the focus.

While it is recognised on p 8 of the Issues paper that, "[e]ven data classified as personal data has differing levels of sensitivities – health data is more sensitive than data about your personal preference for transport to work – and must be dealt with flexibly in terms of facilitating its access", what is important to note is that the very issue of what is personal data is one that Government and Governmental entitles have yet to fully understand.

There is genuine concern also that once data is released and shared, that it may ultimately fall into the wrong hands. The recent Facebook data breach has highlighted to many Australians how far and wide their data may be shared without their knowledge or consent.[6] This has raised the level of awareness of data safety and security and concern of the general population as to how their data is being used. The genuine lack of trust in the way personal data may be secured and managed has also been highlighted by the level of public debate that has surrounded the introduction of recent My Health Record.[7]

'The purpose test'

As is highlighted on p 6 of the Issues Paper, it is imperative that data is shared for the right purposes. The proposed purpose test under the DS&R Act, as set out on p 8, will allow data sharing and release for specific purposes only such as to:

1. **inform government policy making**, which could include understanding cross-portfolio impacts, identifying trends, modelling policy interventions, assessing broader system trends and evaluation to inform future policy
2. **support the efficient delivery of government services or government operations,** which could include evaluating existing programs, modelling program interventions, targeting programs based

6 https://www.straitstimes.com/asia/australianz/facebook-faces-australia-data-breach-compensation-claim

7 https://www.smh.com.au/technology/australians-are-rightly-questioning-my-health-record-says-privacy-commissioner-20180730-p4zui3.html

on user needs, improving services such as by pre-filling forms, administering or enforcing compliance requirements

3. **assist in the implementation and assessment of government policy**, which could include evaluation of policy and programs, analysis of policy and programs by integrating with additional data or information (e.g. analysis by location)
4. **research and development with clear and direct public benefits**, which could include research by government on a particular topic (unrelated to existing policy or programs), cross-disciplinary research, work by research institutions and academics.

While this proposed test is similar to the test in the New South Wales, Victoria and South Australia schemes of Data Sharing and Release, what seems apparent is the potential uncertainties arising from the breadth of the purposes as they are currently stated. For example, what is meant by "research and development with clear and direct public benefits"?

Perhaps, clearer indications of what category of uses would not be permitted under the DS&R Act would assist the Australian public to place more trust in the security and the safety of the data that Government currently hold and that may be potentially shared and released.

Governments and their entities have long collected data and have used that data, unhindered, for their own internal purposes. The general fear of many Australians is that with new sophisticated data analytics, combined with location data, any Governmental data that is released and shared may ultimately end up in third party hands and be used for purposes that they would not have ever considered.

The general tenor and approach taken by the GDPR is that the onus is put upon data aggregators to account for their actions in relation to the data they collect and hold and to ensure full and transparent disclosure of the uses that is being made of the data they hold.

The GDPR empowers individuals by placing them at the centre of data protection and we would suggest that this approach of protecting individual's data rights, first and foremost, should play a more significant role in the proposed data sharing and release governance framework.