DATA AVAILABILITY AND TRANSPARENCY BILL 2020: EXPOSURE DRAFT

ACCREDITATION FRAMEWORK: DISCUSSION PAPER

SEPTEMBER 2020

This is a submission by members of the Australian and New Zealand Real-World Data Network (RADiANT). RADiANT is a researcher-led and researcher-focussed scientific network established to promote policy that supports good practice in the access to and use of real-world data for research. The Network was initiated in response to the findings and recommendations of the Australian Government Productivity Commission into Data Availability and Use in March 2017. It consists of more than 250 members from the health, medical and other research community and has national coverage across the academic, medical research institute and the not-for-profit research sector.

RADiANT members have substantial first-hand experience with the current burdensome, inconsistent and non-transparent approaches to accessing Commonwealth public sector data. We are grateful for the opportunity to comment on this principles-based legislation from the perspective of the Australian research community.

OVERVIEW FEEDBACK

This legislation has the potential to streamline and increase the safe sharing of public sector data through processes that improve transparency, consistency and accountability as well as the re-use of information.

Importantly for the research community, the legislation will enable:

1. A risk-based approach to data sharing underpinned by the internationally accepted five-safes principles
2. The over-riding of secrecy provisions[1] in Commonwealth legislation that currently prevent the sharing of many high-value Commonwealth public sector data
3. The sharing of Commonwealth public sector data with jurisdictional agencies (currently only jurisdictional data can be shared with the Commonwealth, not the other way around)
4. The creation of high-value enduring National Interest Datasets, as recommended by the Productivity Commission's 2015 Report on Data Availability and Use

We believe that the following aspects of the legislation present potential to strengthen and streamline sharing of Commonwealth data for research purposes:

1. The accreditation of Data Service Providers, organisations and users
2. The public listing of accredited Data Service Providers, organisations and users
3. The standardised application of safe data sharing principles
4. The use of standardised Data Sharing Agreements

[1] Secrecy provisions impose secrecy or confidentiality obligations on individuals or bodies in respect of various kinds of information created, collected or received by the Commonwealth about its citizens.

5. The re-use of information submitted for accreditation about organisations, users and projects, including by other jurisdictions
6. The layering of safeguards, as long as they are tailored to the risk, and not imposed in a one-size-fits all approach
7. Recognition of the importance of Aboriginal and Torres Strait Islander people's data sovereignty
8. Independent oversight of the scheme
9. Complaints and redress processes

We understand that the legislation enables data custodians to designate an Accredited Data Service Provider (ADSP) or another Data Custodian to make decisions on their behalf regarding data sharing. We also understand that the review of project modifications (e.g. adding a dataset) may also be delegated to a single ADSP or DC. If this is widely taken up, it would be a huge advance for Australian health and medical research. We believe that such delegations may prove essential given the likely number of requests, and the likelihood that some data custodians may not be able to service the requests in a timely manner. All processes must be scalable. The ADSPs must also be independent; currently some data custodian agencies also conduct research, or compete with research organisations for funding, and this can be problematic. Given the size of the Australian research community and the need for individuals to perform multiple roles, all potential COIs must be transparent.

We also understand that the implementation of new legislation will impart an unavoidable short-term burden on organisations and users, as well as data custodians, as they transition to the new model and system(s).

However, complexity arises from the need for the legislation to cover data sharing purposes with very different data privacy concerns. Specifically, government service delivery requires the sharing of identifiable data, while research and development requires the sharing of data with overt identifiers removed. For research requiring data linkage, identifiable data are shared only with an accredited third party for the purpose of data linkage and according to established privacy-protecting processes.

With this in mind, we raise the following major concerns and welcome the opportunity to work with the ONDC to see if they can be resolved so the Bill can effectively enable streamlined public good research:

1. It is not at all clear how any existing arrangements regarding sharing of Commonwealth data for research purposes can be maintained noting the scope of the Act (Section 8 of the Bill). We note that Section 8 of the Bill lists two of the circumstances under which the Act applies as "*(f) the sharing is done to enable analysis for statistical purposes*;" and "*(g) the data shared is statistical information*." The Bill does not define the term "statistical" and thus the scope of the data. Arguably, virtually all research that uses data involves "analysis for statistical purposes", while "statistical information" would be considered to be the output of such analysis and would include model outputs as well as aggregated and tabular data.

   Although we understand from the Exposure Draft Consultation paper that the Bill will offer an alternative, rather than mandatory pathway, we are concerned that data custodian agencies will feel compelled to apply it to all sharing of Commonwealth data for research purposes, whether this be unit record data or aggregate data. This will add layers of complexity relating to developing data sharing agreements and accreditation of organisations and users, while research use is also still subject to a parallel process of research ethics and governance. Rather than streamlining sharing of data for research purposes, there is potential for the Act to increase costs and complexity, with these extra burdens being borne by the research community. Some of our members have reported that this is already occurring; this will result in delayed insights to policy and practice, increased burden and costs to researchers, and decreased research competitiveness for funding. We would welcome the opportunity to work with the ONDC on the operational definitions.

2. Section 16 (1) of the Bill states *"The project principle is that data is shared for an appropriate project or program of work".* There is lack of clarity about what constitutes a 'project' or 'program'. This has marked implications for the workload of organisations, users, data custodians and the ONDC. We would welcome the opportunity to work with the ONDC on the operational definitions.

3. Section 16(1) of the Bill specifies "*any applicable processes relating to ethics are observed*". There is lack of clarity about which projects or programs need ethics approval, and if it is needed, whether it is required before or after data custodian approval. We understand the lack of clarity probably stems from the fact that ethical approval is not required for government service delivery, while it is usually required for human research. However, many of the information requirements and issues currently considered by Human Research Ethics Committees and institutions (research governance) overlap with those considered under the accreditation framework, the application of the data sharing principles, and the development of the data sharing agreement. This legislation represents a game-changing opportunity to reduce the duplication and complexity of these processes for public good research that re-uses public sector data. Such reform was a key recommendation of the Productivity Commission.

   We note duplication in the requirement for both ethical and data custodian review, and the current co-dependency of data custodian approval on ethical approval, and ethical approval on data custodian approval, by some but not all data custodians and ethics committees. We strongly support the intention (ONDC, verbal communication) that the legislation provides the legal authority for single review. It should be noted that legislation and regulation in other countries (e.g. USA Federal Policy for the Protection of Human Subjects [45 CFR 46]), specify that certain categories of research use of existing datasets (e.g. research using data that does not include personal information) is exempt from a requirement for ethics review. Australia has more restrictive policies regarding exemption from ethics review than the United Kingdom, the United States or the Netherlands, which reduces our research productivity and global competitiveness (Scott AM, Kolstoe S, Ploem MC et al. Exempting low-risk health and medical research from ethics reviews: comparing Australia, the United Kingdom, the United States and the Netherlands. *Health Res Policy Sys* 2020; 18: 11. https://doi.org/10.1186/s12961-019-0520-4). We would argue that Australia should follow the Netherlands by explicitly stipulating that exposing participants to treatment, or requiring them to follow behavioural rules, are the determinants of whether research requires ethics review, and explicitly designating an exemption for secondary use of data that does not include personal identifiers.

4. The accreditation framework is unnecessarily excessive for researchers seeking to access public sector data that does not include personal identifiers in controlled environments (safe havens), for example linked health record data accessible only in accredited secure remote access computing environments such as the Sax Institute Secure Unified Research Environment (SURE), UNSW e-Research Institutional Cloud Architecture (ERICA) or AIHW Secure Remote Access Environment (SRAE). Researchers will never have access to personal information as defined by The Privacy Act, so this adds a new administrative burden. This process is overlaid and does not appear to replace existing administrative requirements, so it does not meet the overarching aims of the Bill.

5. The term “Data minimisation” does not appear in the Bill, only the Exposure Draft Consultation Paper. There is scope for this principle to be open to interpretation, creating a complex approvals and data supply process if there is variation between data custodians for a given project. How will this be standardised in practice, since a one-size-fits all approach is also undesirable? Furthermore, it is not possible or desirable to provide justification for the requirement for each individual variable for projects that apply machine learning techniques to very large datasets. A more practical approach may be to specify certain sensitive variables (e.g. small geographies, Indigenous status) as requiring individual justification, with all others being supplied by default for projects using data that does not include personal identifiers accessed through a ‘safe haven’.

FEEDBACK ON THE EXPOSURE DRAFT CONSULTATION PAPER

Section 1.3 states *“The Bill will not undermine or invalidate existing data sharing arrangements that are working successfully for all participants. Rather, the Bill will provide an alternative pathway to share data where it is currently prevented by secrecy provisions or where it simplifies existing pathways.”* As noted above, it is not at all clear how existing arrangements regarding sharing of Commonwealth data for research purposes can be maintained noting the scope of the Act (Section 8 of the Bill).

Section 2.2. My Health Records have been singled out for exclusion without justification. In the interests of transparency, we suggest a justification is given.

Section 2.2. With respect to the following passage on page 16: “*The National Data Commissioner will not be able to overturn a decision not to share data as there is no duty to share using the Bill. However, as part of their annual reporting, the Commissioner could seek to identify reasons for such denials and improve guidance to resolve any uncertainty leading to those decisions*.” We recommend that the justification for these decisions be made public, so that there is full accountability. This will build public trust and confidence in the system. This would be essential information to help guide community discussion regarding what constitutes “public benefit” and “public interest”, acknowledging this may change over time. Not making the decisions public would be at odds with the following passage on page 18: “*The transparency requirements in the Bill support scrutiny of the scheme by the public and the National Data Commissioner… <u>This transparency promotes accountability across the scheme</u>, <u>putting the onus on decision-makers to demonstrate to the public</u> how the public interest is served by the sharing*”. And the following passage on page 19: “*Such a process inevitably requires subjective judgements to be made, and the scheme requires custodians to apply <u>and demonstrate</u> rigour in this process*”.

Section 4.1. We agree that it will take time for participants to confidently administer and use the new scheme. During this phase, unwarranted precedents may be set, in terms of both project approvals and project rejections. What steps will be taken to minimise such outcomes?

Section 2.2.6. RADiANT members would like to see more detail about the exact public transparency measures the Bill will deliver, including the reports to be tabled in Parliament. We recommend this include performance indicators about the number of project and data approvals, the time taken to grant project/program approval, and details regarding rejected requests (number, data collection, reason).

FEEDBACK ON THE ACCREDITATION FRAMEWORK DISCUSSION PAPER

The model is essentially accreditation to apply for a future Data Sharing Agreement, i.e. a new layer of regulation. Any subsequent Data Sharing Agreement may stipulate further controls for a specific project. No implementation details are provided, but future co-design workshops are planned.

Page 1: Similar to the Exposure Draft Consultation paper, this Discussion Paper states: "*This data sharing scheme is an alternative pathway to share public sector data where it is currently prevented by secrecy provisions or where arrangements are burdensome and complex. It does not replace or impact existing data sharing arrangements which are working successfully.*" Again, we have concerns regarding how existing models for sharing data for research purposes can or will be maintained noting the scope of the Bill, and believe that there is strong potential for the Act to increase costs and complexity, with these extra burdens being borne by the research community.

Page 1: "*Our proposed approach takes into account the views we have heard so far and has been designed to: provide a consistent national framework that states and territories could also leverage in future.*" We are interested in greater transparency about the incentives and barriers for state and territory agencies to adopt this approach so that a nationally consistent framework can be realised. We recommend the development of a single, national infrastructure to support the application and approvals processes.

Page 3: "*Data Custodians will have access to information about accredited entities and their data capability, before deciding whether and how to share data.*" Health research applications typically involve multiple data custodians. Will data custodian approval be sought in parallel or sequentially? Will data custodian feedback be visible to applicants and other data custodians? This framework is not a reform to the current processes without widespread uptake of delegated authority to a single entity. There is an emphasis on building a framework and processes that are "robust and streamlined", yet there is no evident streamlining. It mostly describes the collection of information about users and organisations, not what will be done with this information, nor reforms to the wasteful and complex approval processes.

Page 7: RADiANT supports the inclusion of non-Australian citizens and non-permanent residents as accredited users. We would like greater clarity about the accreditation of postgraduate students who may not have a substantive curriculum vitae. Will the credentials of the primary supervisor be taken into account in such cases?

RESPONSES TO ONDC QUESTIONS

PROPOSED FRAMEWORK

1. **What is considered to be an appropriate level of Australian ownership for an organisation to be eligible for accreditation?**

   No comment.

2. **Should individuals acting on behalf of an Accredited Data Service Provider be accredited individually? If so, what might be appropriate arrangements?**

   This would add another process to what already occurs and is probably not required, unless there are major shortcomings with existing processes. If this is necessary, it may be worthwhile considering having different types of accreditation for different entities/roles. For example, an ADSP that performs record linkage will ensure that its staff are qualified to undertake those tasks.

3. **Are there circumstances when it should be mandatory to use an Accredited Data Service Provider for a data sharing project?**

   Yes.

4. **What would those circumstances be?**

   Whenever identifiable information is to be shared with an Accredited Data Service Provider (i.e. for the purpose of data linkage for a research project). It may be worthwhile considering requiring ADSP involvement when a custodian wishes to share their data but does not have the capacity or capability to do so.

ACCREDITATION CRITERIA

5. **Are there elements of data capability that should be given more or less weight in the accreditation process, i.e. making elements mandatory or optional?**

   Yes, there are several information elements that should be optional for users (i.e. researchers) who will only ever access shared public sector data in a secure, remote analysis environment. In these circumstances, the user will only hold analytical output that has been vetted prior to release.

   In terms of technical skills and capability, it would be onerous and of low value for users to provide professional references.

   For organisations (or departments/centres/units) applying for accreditation who will only access data via an ADSP or safe haven, the requirement to provide information on organisational IT and security infrastructure should be waived.

   As noted above, some exemptions may be necessary for postgraduate students seeking to become an accredited user. In such cases, it may not be necessary to demonstrate experience, rather, the applicant could describe the oversight, training and mentorship to be provided by a suitably qualified and accredited individual (their supervisor).

6. **What elements would be most useful to Data Custodians to support their decision-making process when considering sharing and access to data?**

   Whether the accredited user was going to access the data in a 'safe haven'.

   Custodians would likely wish to ensure that the data is interpreted 'correctly'. It may therefore be helpful to capture how the applicant plans to do this (e.g. partnership / consultation with subject matter experts [particularly for datasets being accessed for the first time], or historical experience with the requested data). We do not think this needs to be a requirement, however, it may lead to more favourable and/or timely assessment of applications.

7. **Should the accreditation process recognise other frameworks, standards or processes that have assessed an element of data capability? If so what standards/processes might be appropriate to recognise?**

   No.

8. **Are there any elements of data capability that should be captured in order to understand an accredited entity's ability to keep data safe?**

No comment.

ACCREDITATION PROCESS

9. **What is a reasonable period of time to assess an application?**

One month. It may be possible to consider different assessment periods for different types of application (e.g. one month for organisations, less time for an individual within an accredited organisation).

If an applicant is required to undertake training in order for their application to be approved, then the training needs to be offered frequently, or on an on-demand basis (e.g. online).

10. **Are there further ways we can streamline the accreditation process?**

We recommend the ONDC develop case study examples of what would constitute "sufficient evidence" for accreditation.

Many users will seek to perform similar projects, in the same setting (secure, remote analysis environment), with the same or similar data and outputs (including independent vetting of outputs). We therefore strongly recommend the accreditation go further and broadly accredit users ("trusted users"), as per Productivity Commission recommendations) with respect to National Interest Datasets, certain types of (i) projects, (ii) settings, (iii) data and (iv) outputs, as well as prescribed combinations of (i-iv). Otherwise, each data custodian will be repeating the efforts of the ONDC in reviewing the qualifications, experience, governance and administration of Accredited Users and organisations. This process is duplicative and wasteful. What is preventing such accreditation from being centralised and standardised, following the development of consensus-based guidelines?

The accreditation process as currently designed does not result in a "trusted user", rather a "licensed user", and thus does not "build trust" in the system. User credentials will still be reviewed on a project-by-project basis by multiple decision makers.

There is potential for significant delays with respect to perceptions of foreign interference. How would this be dealt with in practice considering, for example, the multitude of arrangements that universities and other organisations have with overseas-based organisations?

11. **Do the timeframes to renew accreditation, every 5 years for Accredited Data Service Providers and every 3 years for Accredited Users, seem reasonable?**

Yes.

OTHER MATTERS

12. Is it appropriate to notify parties to Data Sharing Agreements of an accredited entity's suspension?

No, because it may unfairly impact future requests for Data Sharing Agreements by an Organisation or User, even when the Commissioner rules that there was no foundation to the suspected breach.

13. **Is there any information that must, or must not, be made publicly available through the registers of accredited entities?**

No.

14. **Is there any information that should be made available to Data Custodians through the registers of accredited entities?**

    All information provided during the accreditation process should be available to Data Custodians.

15. **Is charging a fee for accreditation, such as a renewal fee, reasonable?**

    Given the available details of the renewal process, there does not appear to be strong justification for charging a fee for renewal of accreditation.

    We agree with the statement that fees could be a barrier for some organisations or individuals to apply for accreditation, which would reduce Australian research productivity.

TRANSITION ARRANGEMENTS

The proposal for Accredited Integrating Authorities to transition to Accredited Data Service Providers appears worthwhile. Furthermore, allowing these entities to perform data sharing on behalf of a data custodian is a transformational step that will reduce the time taken for data sharing, if these entities are appropriately funded to perform this service.