Dear Dr Gould,

Data Sharing and Release Legislation Consultations

The University of Sydney welcomes the opportunity to provide feedback on the New Australian Government Data Sharing and Release Legislation, Issues Paper, released for consultation on 4 July 2018. Our attached submission seeks to supplement the feedback our representatives provided at the helpful face-to-face consultations sessions you facilitated in Sydney on Friday 27 July 2018.

As a public educational and research institution that exists to deliver outcomes for the benefit of Australia and the wider world, we have a deep and longstanding interest in the Australian Government’s efforts to improve the availability of data it holds for research and other purposes that benefit the community. We engaged actively with the Productivity Commission’s Data Availability and Use inquiry in 2016 and subsequently participated in discussions between the Department and university sector about the Commission’s report prior to the Government releasing its response in May 2018.

Our interest in improving the availability and use of data held by Commonwealth entities is driven by two key factors. First, various Commonwealth agencies hold data about our education, research and associated activities that will be within the scope of the proposed new legislation. We are therefore keen to understand the implications of the proposed new legislation for our operations as an extensive provider of data (sometimes sensitive) to many Government agencies. This includes being part of designing the processes by which data sets relating to universities, their staff, students and alumni will be made more accessible. Second, as a large research organisation with many hundreds, if not thousands, of researchers seeking access to data held by Commonwealth entities each year, we strongly support the policy goals the Government’s proposed reforms are designed to achieve. We are keen, however, to ensure the legislation and processes the Bill establishes do indeed serve to streamline, rather than stifle, the release of Commonwealth data sets for research and other public interest purposes.

As a major research and research training organisation and custodian of thousands of data sets (many arising from research that has been supported financially by the Australian Government) we are grappling with similar issues and challenges to those the Government faces on an even larger scale. Like the Government, we are committed to maximising open access to the publications of our researchers, as well as to the data they use or generate for their research. Our efforts to maximise data access occur in accordance with the open access requirements of the Australian Research Council (ARC) and National Health and Medical Research Council (NHMRC) and with an eye to developments and leading approaches internationally. For many years we have been investing heavily to enhance our data infrastructure, analytical capability, coordination, support services, policies and processes to manage risk and ensure that our researchers and research students have access to high quality facilities, training and support regarding all aspects of research data management and responsible use.
Our submission provides feedback on all sections of the Issues Paper, raising key issues for consideration including the following six.

1. Governance, consultation and involvement of the research community

The establishment of the National Data Advisory Council (NDAC) is supported. However, we have reservations about how effective it will be, particularly regarding matters of research, given the multifaceted and complex nature of the issues involved. We are concerned that no single academic will be able to advise adequately on all issues that are relevant to the effective conduct of research, nor effectively represent the views of the research community. The NDAC would therefore benefit from having its membership expanded to include ex-officio members nominated by their peak research sector or professional bodies. Alternatively, or in addition, the NDAC and Commissioner would benefit from the establishment of a ‘Research Steering Group’ comprising a mix of representatives of peak sector research bodies experts in all relevant aspects of research involving data.

2. Policy objectives and desired outcomes

The stated objects of the Bill and its ‘purpose tests’ are heavily focused on the purposes for which data will be used to improve Government processes and policies. As the Government’s response to the Productivity Commission’s report on Data Availability and Use stated: ‘The reformed national data system will greatly improve Australia’s ability to capture the social and economic benefits from existing data.’ However, at Page 6 the Issues Paper states that the purpose of the Bill is only ‘to streamline the process for sharing public sector data and improve data safeguards across the public service.’ The Bill and purpose tests would benefit from the inclusion of a much clearer statement of the policy objective in terms of the public benefit outcomes the Government is seeking, with each of the permitted data usage ‘purposes’ linked clearly to the realisation of this overarching policy goal.

3. The definition of ‘public benefit’

The range of activities covered by the fourth proposed purpose test ‘Research and development with clear and direct public benefits’ is very much open to interpretation and potentially highly restrictive in scope. The term “R&D” is well understood and defined in other existing Commonwealth Acts and guidelines. However, given the inherently uncertain nature of some research, the term ‘public benefit’ will need to be clearly and broadly defined, and applied consistently for all data release decisions. Otherwise, there is a real risk that the release of some data sets for research purposes will be restricted, resulting in the full potential benefits of the data’s release never being realised. Further thought is also needed in relation to the treatment of activities, which while not research, may nevertheless have the potential to delivery substantial public benefits.

4. Accredited Data Authorities and ‘Trusted Users’

The Productivity Commission recommended that several Accredited Data Authorities (ADAs) should be established to share the expected high volume of requests for access to datasets, and provision of specialist data services (such as data curating and linkage). Many researchers currently wait unreasonably long periods (sometimes measured in years) to access linked datasets provided by existing Integrating Authorities. Every effort should be made to improve this situation and we agree that the number of ADAs authorised to link and release data should be increased.

The Issues Paper is unclear about whether the Department envisages ‘Trusted Users’ being individuals, research organisations such as universities, or a mixture of both. In line with the approach recommended by the Productivity Commission, ‘Trusted User’ accreditation should generally apply to research institutions, which would then be responsible for providing the support and resources to their researchers to become ‘safe people’ and to abide by the applicable legislation, codes and data sharing agreements.
Getting the system for ADAs and Trusted Users right will be critical to the success of the reforms in terms of streamlining researchers’ access to data. Given the high stakes and complex range of issues that will need to be worked through to design and administer the research aspects of the new framework, it is critical that the Department engages with the Australian university sector and broader research community to agree the most cost-effective and efficient approach for the operation of the new arrangements for the release of data.

5. National Interest Data Sets

The Issues Paper appears silent on the question of how National Interest Datasets will be treated under the proposed Bill. In its final report, the Productivity Commission recommended the identification (through expert, political and community consultation) of strategically important National Interest Datasets (NIDs), and appropriate infrastructure and funding to build and maintain them. These datasets, linking government data from various Federal, State and Territory sources, would form valuable resources for governments and researchers who could use them to produce insights of potentially enormous public benefit across diverse areas of policy. We encourage the Department to ensure NIDs are not forgotten as the detail of the legislation is developed.

6. Review of data release and other decisions made under the Act

The data sharing process flowchart (Issues Paper page 13) is helpful in showing the decision-making framework for requests to release data. However, a critical review process has been omitted. There should be provision built in to review applications for data release that are denied by a data custodian, ADA or National Data Commissioner (NDC), or where unreasonable terms are applied in respect to the release of data. Data custodians and ADAs should be encouraged to establish appropriate internal merits review processes to allow for a quick and inexpensive means of re-examining any primary decisions to deny access to data, where an applicant believes an error has been made. It is important that there is effective and timely recourse should this occur. The NDC should also have appropriate powers of review in respect of decisions made by data custodians and ADAs to ensure the legislation is applied consistently and fairly. Once the role and powers of the NDC are known, clarity will also be required about the processes for review of data release and other decisions he or she may make under the Act.

We trust our feedback is helpful and look forward to continuing to work with your office and other stakeholders to help ensure this critically important piece of legislation achieves its objectives.

Yours sincerely,

(Signature removed for electronic distribution)

Professor Duncan Ivison
Deputy Vice Chancellor, Research

Attachment University of Sydney submission to the New Australian Government Data Sharing and Release Legislation, Issues Paper for Consultation, released 4 July 2018
Overview

The proposed new legislation outlined in the Department of the Prime Minister and Cabinet’s Issues Paper for Consultation document released 4 July 2018 covers the sharing of all data collected or developed by Australian Government entities and Commonwealth companies as defined under the Public Governance, Performance and Accountability Act 2013. We note that the recommendations made by the Productivity Commission in its Inquiry into Data Availability and Use addressed broader data issues, including steps to improve access to privately-held (commercial) datasets and the rights of individuals to access those data.

While this consultation process is limited to addressing how the proposed Data Sharing and Release Bill (DS&R Bill) concerning Australian Government datasets can be drafted and implemented effectively to stimulate and enhance research opportunities and other outcomes in the public interest, we are also keen to ensure that the work Treasury is leading to develop the Consumer Data Right complements these objectives. It is vital that both elements of the Government’s planned response to the Productivity Commission's report work effectively and consistently together toward the Government's policy objective of ‘greatly assisting consumers, researchers, government agencies and industry to better understand the world we live in and to make sound investment decisions based on evidence.’

We are confident that this can be achieved if the new legislative framework is developed and implemented in close consultation and collaboration with stakeholders including universities and other public and private sector research organisations.

Our response to the Issues Paper's consultation questions

Rather than seek to answer each of the 43 questions raised in the Issues Paper, we have below provided comments on each of the sections (i.e. groups of questions) presented in that document.

---

Issues Paper Section 2. Key Principles of the Data Sharing and Release Bill

Consultation questions

1. Are these the correct factors to be taken into account and to guide the legislative development?
2. What else should the Government take into consideration when designing the legislation?

• While the factors discussed in the Issues Paper are valid, they will only promote better sharing of public sector data if the roles and responsibilities of the ‘end users’ of the newly accessible data are properly identified, with their requirements accommodated. For the research sector, there is the very real danger that a new framework will result in increased layers of bureaucracy that will stymie, rather than facilitate, the Government’s goal of making the data it holds more readily accessible for researchers.

• The creation of the new office of a National Data Commissioner (NDC) to oversee the framework is commendable, but we have reservations about how effective and efficient the NDC can be given the Office’s extensive range of duties and responsibilities, limited resources and the potentially large number of stakeholders involved in each decision (e.g. OAIC, ABS, ADAs, government departments, end users).

• The establishment of a National Data Advisory Council (NDAC) to advise the NDC is an important part of the framework; intended to represent views, advice and expertise on a range of data issues from senior academic, industry and privacy representatives. However, we have reservations about how effective the NDAC will be, the level of authority and influence it will carry and the value that will be placed on its advice. The terms of reference for this group should be made publicly available as soon as possible. We are concerned that we (and other universities) were not officially notified directly of the EoI process and timeframes for membership of the Council and had an extremely limited timeframe to consult with our academic community and ask potential representatives to self-nominate. We suspect that many other research institutions missed this opportunity entirely. The recruitment process for the appointment of Council members is yet to be made available publicly. The lack of transparency around this process seems to be at odds with the aims of the proposed legislation.

• Given the critical and complex nature of the issues at stake in relation to access by public and private sector research organisations to Australian Government data sets, the proposed NDAC
would benefit from having ex-officio expert members nominated by their peak sector or professional bodies. Alternatively, or in addition, the NDAC would benefit from the establishment (in the development and early stages of implementing the new Act at least) of a dedicated ‘Research Steering Group’ or ‘Expert Advisory Group’ comprising experts in the efficient storage, management, sharing, legal and ethical use of publicly held data sets for research and analysis purposes as well as research administrators familiar with best practice approaches to data release by Australian and international government agencies.

**Issues Paper Section 3. Scope of the Data Sharing and Release Legislation**

**Consultation questions**

3. Should the scope be broader or narrower?
4. Are there entities that should be included or excluded from scope? How would this be justified?
5. Should any specific categories of data be specifically out of scope? How would this be justified?
6. Should exemptions, for example for national security and law enforcement, occur at the organisational level or for specific data categories?
7. Are there instances where existing secrecy provisions should prevail?

- The Issues Paper suggests that the proposed legislation will apply to all Commonwealth entities and companies. The Act will need to apply consistently to all statutory public research organisations (whether established by Commonwealth, State or Territory legislation) to ensure that a level-playing field for data governance and access is preserved for all Australian research organisations. For example, while most Australian universities are established by Acts of State parliaments, the Australian National University (ANU) was created by a Commonwealth Act and so would appear to be within scope as a ‘Commonwealth entity’. Moreover, numerous Commonwealth research agencies are users of data held by other Commonwealth entities for research purposes as well as the custodians of data sets, for instance:
  - Australian Antarctic Division (AAD)
  - Australian Astronomical Observatory (AAO)
  - Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS)
  - Australian Nuclear Science and Technology Organisation (ANSTO)
  - Commonwealth Scientific and Industrial Research Organisation (CSIRO)
  - Defence Science and Technology Organisation (DSTO)
  - Geoscience Australia (GA)
  - The Australian Institute of Marine Science (AIMS)
• Consideration will need to be given to ensuring the Act applies consistently to all Australian universities regardless of whether they are established by Commonwealth, State and Territory Acts of Parliament, as both a provider of (in some cases sensitive) data to the Commonwealth about their activities and as an accessor for research purposes of data sets held by a wide range of Commonwealth entities. Common mechanisms will also be required to ensure that the data access and usage requirements of the Act, at least, apply consistently to Commonwealth publicly funded research agencies as compared to similar research institutions established by State or Territory Governments.

• Designating specific categories of data as ‘out of scope’ should be avoided wherever possible. Datasets not containing sensitive information should be made broadly accessible; access to datasets containing sensitive data should be assessed on a case-by-case basis, with data custodians/ADAs de-identifying data before release. High security datasets (classified) would not be eligible for release.

• The Productivity Commission recommended the nomination and establishment of National Interest Datasets (NIDs) that would be of high value to many stakeholders, including researchers, and potentially would be managed through a different framework. The scope of the legislation should be expanded to include NIDs.

Issues Paper Section 4. Streamlining Data Sharing and Release

Consultation questions about the purpose test
8. Do you agree with the stated purposes for sharing data?
9. Are there any gaps in the purpose test that would limit the benefits of public sector data use and reuse?
10. What further detail could be included in the purpose test?
11. Should data be shared for other purposes? If so, what are those purposes?
12. Should there be scope to share data for broader, system-wide purposes?
13. Should the purpose test allow the sharing of data to administer or enforce compliance requirements?
• The stated objects of the ‘purpose test’ are focused mostly on the purposes for which data will be used around improvements for government processes and policies. They generally lack a higher vision of the potential benefits for the community and economy of the proposed maximisation of external access to Commonwealth-held data sets subject to appropriate safeguards. As the Government’s response to the Productivity Commission’s report on Data Availability and Use stated: ‘The reformed national data system will greatly improve Australia’s ability to capture the social and economic benefits from existing data.’ The Issues Paper states at page 6 that the purpose of the DS&R Bill ‘will be to streamline the process for sharing public sector data and improve data safeguards across the public service.’ The purpose test would benefit from the inclusion of a clear statement of the overarching objective of the legislation in terms of outcomes the Government is seeking, with each of the permitted data usage ‘purposes’ then linked clearly to this higher-level policy objective.

• The fourth stated purpose ‘Research and development with clear and direct public benefits’ is very much open to interpretation and again potentially restrictive in scope. While the term “R&D” is well understood and defined in other existing Commonwealth Acts and guidelines, the terms ‘public benefit’ will need to be clearly defined in the Act, and that definition must recognise that the immediate ‘public benefit’ of many research outputs is often not apparent at the time that access to data is requested. The inherently uncertain nature of the incremental/iterative research processes to build on previous knowledge (using data held by Commonwealth entities) to deliver innovations and benefits needs to be accommodated in the purpose test. There is a real danger that if the ‘public benefit’ test is applied too narrowly, data releases for research purposes will be restricted, resulting in the full potential benefits of the data’s release never being realised.

• Consistency in the application of the purpose test needs to be ensured. Different data custodians and ADAs must all work from the same set of standards when determining whether requests for data pass the purpose test. The application of the test to determine if research is ‘R&D’ for ‘public benefit’ must satisfy a prescribed set of criteria that are universal and defensible, with an appropriate framework for internal and external review of any primary decision to refuse a request for access. We suggest strongly that the Advisory Council and the wider research community be engaged to determine what these criteria should be. Given application of such criteria involves discretion on the part of the primary decision maker, we also suggest the legislation and/or

---

2 See for example the OECD Frascati Manual and the Australian Government’s Higher Education Research Data Collection Specifications for the internationally agreed definition of R&D.
accompanying explanatory materials would sensibly include guidance and support to assist those applying the criteria, including examples of relevant considerations.

- The Productivity Commission report emphasises the power of improved data access to underpin new economic activity by enabling the development of ‘new products and services that transform everyday life, drive efficiency and safety, create productivity gains and allow better decision-making.’ While making public sector data more readily accessible for research and development purposes will help unlock such innovations, data may be accessed and used for purposes that may not be research and development, but which nevertheless may deliver substantial public benefits. For example, a private company may access data from a Commonwealth agency and analyse or present it to consumers in a novel way that makes the data more accessible or useful to individuals. The outcome may be a mix of private and public benefits through the delivery of new services. The ‘R&D’ and ‘public benefit’ test should not stand in the way of this type of activity, which has the potential to deliver substantial economic and social benefits through the creation of new services, industries and jobs based on improved access to Commonwealth-held data sets.

- The process flowchart (page 13) is helpful in showing the decision-making framework for requests to release data. However, a critical review process has been omitted. There should be provision to review applications for data release that are denied by the data custodian, ADA or NDC or where unreasonable terms are applied in respect to the release of data. Data custodians and ADAs should be encouraged to establish an appropriate internal merits review process to allow for a quick and inexpensive means of re-examining any primary decision to deny access to data, where an applicant believes an error has been made. There is a real possibility that valid requests will be refused if data custodians apply too narrow an interpretation of ‘the public benefit’ test to release requests for research purposes. It is important that there is effective and timely recourse should this occur. For this reason and as further discussed below, the NDC [or another suitable external body] should also have appropriate powers of review in respect of decisions made by data custodians and ADAs to ensure the legislation is applied consistently and fairly.

The National Interest Datasets are not included here, and they should be. In its final report, the Productivity Commission recommended the identification (through expert, political and community

---

3 Productivity Commission, Data Availability and Use, Productivity Commission Inquiry Report, Overview and Recommendations, No.82, 31 March 2017, p.2
consultation) of strategically important National Interest Datasets (NIDs), and appropriate infrastructure and funding to build and maintain them. These datasets, linking government data from various Federal, State and Territory sources, would form valuable resources for governments and researchers who could use them to produce significant insights of enormous public benefit into social, health and economic issues.

**Consultation questions about data safeguards**

14. Is the Five-Safes framework the appropriate mechanism to ensure data is safeguarded?
15. Are there any additional safeguards that should be applied?
16. Are there any instances when the Five-Safes could not be applied?
17. Is the Five-Safes appropriate when data is shared and used for the specific purposes in the purpose test above?
18. How should the responsibility for managing risks be shared in the framework?
19. How would you envisage Five-Safes principles be applied over the life-cycle of data to ensure data safeguards are continually met?
20. Under what circumstances should trusted users be able to access sensitive data?

**Consultation questions about public sector data sharing arrangements**

21. Would this arrangement overcome existing barriers to data sharing and release?
22. Would streamlined and template agreements improve the process?
23. Do you agree that data sharing agreements should be made public by default?
24. What level of detail should be published?
25. What else should a data sharing agreement contain?
26. What other transparency mechanisms could be mandated?

- The Five-Safes Framework is appropriate to use. It is an internationally accepted method for evaluating and managing risks when releasing data. The CSIRO Data 61 De-identification Decision-Making Framework should be used to assist in releasing data containing identifiers, where appropriate to do so.

- Submitting requests for data sharing should be as quick and efficient as possible. Streamlined processes would be welcome, and presumably templates for requests will be provided to applicants. Regular communication between applicant and authorising body of progress throughout the application process will be necessary.
• Transparency of data sharing agreements of government data is to be encouraged. However, the amount of information released about each agreement should be carefully regulated to protect the identity of researchers working on sensitive topics or where research IP may be compromised.

Issues Paper Section 5. Roles and responsibilities within the system

Consultation questions
27. How long should accreditation as an ADA or Trusted user last?
28. What could the criteria for accreditation be?
29. Should there be review rights for accreditation?
30. Should fees be payable to become accredited?
31. Is the Australian Government Charging Framework fit for purpose in this context?

• The Productivity Commission’s recommendations were that several Accredited Data Authorities (ADAs) should be established to share the expected high volume of requests for access to datasets, and provision of specialist data services (such as data curating and linkage). Many researchers currently wait unreasonably long periods (sometimes measured in years) to access linked datasets provided by existing Integrating Authorities (AIFS, ABS and AIHW), particularly when Human Research Ethics Committee approval is required too. Every effort should be made to improve this situation and we agree that the number of ADAs authorised to link and release data should be increased.

• The accreditation of non-government entities as ADAs requires careful consideration. ADAs will be responsible for risk management, governance, compliance and decision-making around release of public data, and could be required to make decisions to release data that may conflict with commercial or research imperatives. For example, it may not be in the interests of a non-government ADA (part of a larger data services entity, perhaps) to release data to a competitor. Should a reasonable request from a competitor be refused, would the requestor have recourse? Would the denial of service by the ADA be evident (i.e. accountable)? Additionally, the public perception that a private-sector ADA would be profiting from services derived from commercialisation of government (i.e publicly-owned) data may be problematic.
We agree that some sort of accreditation for safe data handling is necessary. However, the definition of what constitutes a ‘trusted user’ is unclear. Does this mean that each individual user of government data in a research environment needs to become registered as a trusted user? Or only the Chief Investigator/research group lead for a research project that accesses the data? Or a delegate of the CI/research group head who is responsible for acquiring and managing the data? Or is it the research institution?

In line with the approach recommended by the Productivity Commission, the ‘trusted user’ accreditation should apply to a research institution; the institution would then be responsible for providing the support and resources to its researchers to become ‘safe people’ and to abide by the applicable legislation, codes and data sharing agreements. All researchers who work in an Australian research institution and access any data (wherever it is sourced from) are bound to that institution’s code of conduct, data management and privacy policies. The newly-created NDC would be responsible for accreditation of the ‘trusted user’ research institutions, and the institution would comply with directions from the NDC regarding implementation of the Five-Safes Framework.

Alternatively, if the ‘trusted user’ concept is aimed at individual researchers or research groups, then the costs of compliance for the Government, individuals and research entities are likely to become very high. The Issues Paper states that accreditation would be maintained and audited by the NDC, so an individual wishing to enter a data sharing agreement would need to first apply to the NDC to be accredited (with inbuilt review process should accreditation criteria not be met). What would this process entail, and should the accreditation criteria be linked to the types of data that the applicant wishes to access? (e.g. police checks, security clearances, nationality/visa status, affiliations, specialist training/declaration, access to secure systems for data processing etc.). Perhaps different categories of accreditation would apply for access to different types of data. There are many possible permutations, all of which would add complexity to a bureaucratic process. Presumably, the accreditation process will need to be repeated when the initial accreditation period lapses.

---

• It is not clear whether the NDC will be funded adequately to administer an accreditation system that is efficient and does not add unacceptably long delays or costs to requests for government data access. Accreditation fees should be avoided if possible. It would make more sense for data access fees to be charged (most of which would be recouped from specific research grants).

• We reiterate that given the high stakes and complex range of issues that will need to be addressed in designing and administering the new data access framework for research purposes, it is critical that the Department of Prime Minister and Cabinet engages constructively with the Australian universities and the broader research community to agree on the best approach to design and implementation of the new regime for data sharing and release.

Issues Paper Section 6. National Data Commissioner

Consultation questions

32. Are these the right functions for the National Data Commissioner?
33. What review powers should the National Data Commissioner have?
34. Should the NDC have the power to conduct an investigation into system-wide issues?
35. What other actions could the NDC be able to take?
36. Are there other ways community values and expectations can be captured and addressed?
37. What aspects should be taken into consideration when considering consequences for non-compliance with the DS&R Bill?
38. Should the consequences differ depending on the type of data involved or the type of misuse, e.g. harsher penalties for intentional misuse?
39. Should penalties be strict liabilities?
40. What would be an appropriate penalty for intentional misuse of data?
41. How would responsibility for misuse of data be shared across the data system?
42. To what extent should there be a complaints mechanism and how should it work?
43. Should a complaints mechanism provide for complaints by the public?

• Given the decentralised approach to decision making contemplated by the DS&R Bill, as well as the discretionary nature of tests and frameworks to be applied, there may be substantial variation in how different data release and accreditation requests are assessed and decided. The NDC
should have appropriate powers of review over all key decisions made under the legislation including denials of accreditation of trusted users/institutions, contested refusals by data custodians to access data or disputes regarding the terms which apply to such access. The review processes within the data release framework when accreditation or access requests are denied need to be clearly defined, implemented and observed by data custodians, ADAs and the NDC itself. To ensure a consistent and whole-of-government approach, NDC reviews of data custodian and ADA decisions must necessarily be merit reviews, constituting a fresh consideration of the facts, law and policy aspects of the original decision to determine the correct or preferable decision. Consideration might also be given to whether external review is appropriate for any such decisions.

- Penalties for misuse of data are already covered by existing legislation, and by OAIC. The NDC would be able to revoke any accreditation for serious breaches.