ACADEMY OF
THE SOCIAL SCIENCES
IN AUSTRALIA

The Academy of the Social Sciences in Australia's Response to the New Australian Government Data Sharing and Release Legislation

13 August 2018

Prof Glenn Withers (President)

Prof Diane Gibson (Policy and Advocacy Committee Chair)

Dennis Trewin (Policy and Advocacy Committee Member)

26 Balmain Crescent, Acton ACT 2601

GPO Box 1956, Canberra ACT 2601

P: +61 2 6249 1788

ABN: 59 957 839 703

Table of Contents

1. Introduction

The Academy of the Social Sciences in Australia (ASSA) commends the Australian Government on its commitment to improving Australia's use of data as a key opportunity to substantially enhance national productivity. The importance of improving data use practices extends well beyond questions of national productivity to questions of wellbeing and national progress. Making data more available is an achievable reform and would liberate economy-wide productivity improvements over many years. We welcomed the release of the Australian Government Public Data Policy Statement in 2015, and the recommendations of the Productivity Commission in its *Data Availability and Use Inquiry in 2017*. The proposed new legislation is an important step in reforming and enhancing access to public data in Australia.

Three important issues not raised in the Issues Paper are the need to upgrade the capability in the Australian Government, and more generally, in respect of: (1) the analysis of large and imperfect data sets including linked data sets, (2) the management of these data sets and the design of the underlying systems, and (3) managing the security of these data sets (or at least knowing when to call in the experts).

Although privacy aspects are addressed in several places, we think this needs to be given more attention before the legislation is drafted. This will be scrutinised heavily and it could undermine the effectiveness of the initiative. The most recent Australian Census suffered because the privacy issues were not managed effectively. Another example is the push back on My Health because of

privacy concerns. We strongly support the initiative but we don't want 'the baby thrown out with the bath water' because privacy aspects were not given sufficient attention. The Paper states that 'The Public Policy Statement seeks to ensure the value of public data is maximised'. We would argue that there should be a qualifier such as 'subject to appropriate privacy and data security safeguards' to ensure the message is clear.

The main contributors to this response are Glenn Withers (Fellow and President of the ASSA, and President of the Australian Council of Learned Academies), Diane Gibson (Fellow of ASSA, Chair of ASSA's Policy and Advocacy Committee and former Senior Executive of the Australian Public Service) and Dennis Trewin (Fellow of ASSA, former Australian Statistician and author of international guidelines on researcher access to microdata).

Preamble

The work of government agencies has generated significant administrative data, which is exploited for the purposes of the APS. Research conducted in universities has produced significant data too. Data from each should cohere easily and naturally. This would create what Jim Walter called 'a virtuous circle of benefit both to continuing research and to the policy domain'.

Many kinds of data are used in Australian research, including census data, population and housing data, longitudinal data of individual life courses, data from national surveys, time series data of economic and social phenomena, administrative and business data, and linked data sets. Infrastructure to support this has been established by many agencies, including government agencies such as the Australian Bureau of Statistics, Australian Institute of Health and Welfare, and Department of Social Services, and funding programs such as the Australian Research Council, National Health and Medical Research Council, and National Collaborative Research Infrastructure Strategy.

Despite the work these agencies have done to support this data, it is not presently possible to make this whole qualitatively diverse and quantitatively great range of data available in an open environment. Academic and policy research would require this data's availability at the unit record level, but privacy and confidentiality requirements preclude this.

The data is scattered through the public sector, and among organisations, collecting institutions, and individual researchers and projects. Large amounts of data is hidden in various ways or lays in unstructured forms—for instance, as texts, maps, audio documents, and so on—which make it hard to retrieve. All in all, only a small amount of the data congenial to academic and policy

research are supported by the infrastructure which would make them useful and accessible.

This partly the result of limited and inconsistent funding. Efforts to develop infrastructure have been directed to individual researcher and institution-level priorities, and currently much HASS infrastructure has been project-based and operates at an institutional level. Data infrastructure is uncoordinated and minimally integrated.

This is the context around the question of data sharing. An integrated national data infrastructure should address this by enabling collection, analysis, curation, and the utilisation of data.

Summary: ASSA urges that the importance of developing, supporting and resourcing appropriate data infrastructure in the form of integrated HASS (Humanities, Arts and Social Sciences) data platforms is recognised in the legislation to enhance data sharing and access in Australia.

Response to the Questions Raised in the Issues Paper:

2. Key Principles of the Data Sharing and Release Bill

Questions:

Are these the correct factors to taken into account and to guide the legislative development?

What else should the Government take into consideration when designing the legislation?

Response:

ASSA supports the key principles set out in the consultation document as the correct factors to take into account in guiding the legislative development. We would propose five additional factors be given careful consideration:

a. *The regulatory framework*

An appropriate regulatory framework is essential if the outcomes sought for this project are to be achieved. Models such as the regulatory pyramid have gained considerable traction in social science research, and serve to underpin successful ongoing implementation of complex projects such as this. The regulatory framework for the Act should not be seen as an afterthought, but rather an integral part of the proposal. ASSA believes this is consistent with the principles-based framework outlined for the legislation.

b. *Data Infrastructure*

ASSA has previously lobbied government to support the enhancement of data infrastructure in the social sciences, the humanities and the arts. Data infrastructure of this kind includes a commitment to data comparability, metadata standards, data dictionaries and associated documentation, data management protocols, system inter-operability and enhanced capacity for multiple uses of data once collected. The proposed legislation could be strengthened if the principles associated with enhancement of high quality social science data infrastructure were explicitly included.

c. Improving capability

Related to this is improving capability in data analytic, information management and data security.

d. Ethics approval

The issues paper is silent on the matter of the interface with existing Ethics Committees or the need for ethics approval more generally. While it is noted that Ethics Committees are both more common and more influential in universities and in the health sector, their role in ensuring research integrity has been significant in Australia. We recommend this interface be considered, and the potential role of Ethics Committees in relation for example to the purpose test for data sharing.

e. Commercial data use

The question of access to data by commercial entities, or universities working for or in partnership with commercial entities, is an important consideration. As universities become increasingly entrepreneurial in their focus, the potential to make commercial gain is an important driver. In either case, what are the implications for access to public data, provided for the public good, serving as a profit source for commercial entities? This will be a sensitive question and needs to addressed explicitly. We recommend that these purposes be approached with extreme caution, at least in the initial legislation.

Additionally, the following two elements could also be of relevance in future work:

f. Data linkage

Data linkage is an important resource in building valuable databases from administrative data. It would be useful to include a recognition of the value and the potential risks of data linkage in subsequent work. If data linkage is to be covered by the data sharing arrangements, this is. likely to require specific provisions in the legislation. It is a topic requiring specific consideration as the level of public concern regarding privacy can be escalated.

g. Variegated data systems

Our national data systems vary substantially in their origins and history, and therefore in their quality and capacity to be easily accessed in a data sharing process. Old programs still administered through old computer systems present different challenges to newer and more user-friendly systems.

3. Scope of the Data Sharing and Release Legislation

Questions:

Should the scope be broader or narrower?

Are there entities that should be included or excluded from scope? How would this be justified?

Should any specific categories of data be specifically out of scope? How would this be justified?

Should exemptions, for example for national security and law enforcement, occur at the organisational level or for specific data categories?

Are there instances where existing secrecy provisions should prevail?

Response:

ASSA supports the proposed scope of the legislation, while noting that key elements of the national health and welfare system are delivered by state and territory governments, and that the associated data resides with those entities. While acknowledging current circumstances do not allow the expansion of scope to include state and territory data, we nonetheless urge the relevant authorities to set in places strategies that would enable this expansion of scope in the future.

More generally, it is worth pursuing the sharing of data with States and Territories. However, this will be complex and we suggest it occur after the initial Data Sharing and Release legislation is put in place.

Where national security and law enforcement matters require exemptions from the Act, ASSA supports their application at the specific data category level, rather than at the organisational level. A useful example would be the case of defence, where valuable data is available in health records which does not impact on matters of national security.

Other categories may need to be exempted if the prime purpose of the data collection could be compromised. For example, the reliability of the Census data may be compromised if the public perceive that this data will be shared. The existing legislative arrangements, perhaps with some

sensible modifications, should take precedence over the Data Sharing and Release arrangements. This does not preclude the release of unidentified data or the linking of ABS and other data sets but is should be done under the provisions of the Census and Statistics Act. There may be other similar examples.

4. Streamlining Data Sharing and Release

Questions about the purpose test:

Do you agree with the stated purposes for sharing data?

Are there any gaps in the purpose test that would limit the benefits of public sector data use and reuse?

What further detail could be included in the purpose test?

Should data be shared for other purposes? If so, what are those purposes?

Should there be scope to share data for broader, system-wide purposes?

Should the purpose test allow the sharing of data to administer or enforce compliance requirements?

Response:

ASSA supports the proposed purpose test. It would be helpful if the process for assessing privacy was either more clearly articulated or flagged as an important area for implementation. Ethics Committees may have a role in some cases or independent privacy reviews. Dispute resolution provisions may be needed.

One area that is not explicitly covered, and is valuable, relates to methodological and data infrastructure innovation and testing. For example, the statistical linkage key, SLK 581, commonly used in administrative data collections in health and community services collections, could only be successfully developed and tested because of access to secure national databases.

ASSA recommends that the purpose test should not allow the sharing of data in order to enforce individual or organisational compliance. We do support its use for compliance monitoring purposes

where individuals or organisations are not identified, but research on the scope or nature of non-compliance meets the other elements set out in the purpose test.

There are proposals for the self-management of risks. This could be dangerous if the agencies do not have the capability to assess risks (e.g. information security). They should be encouraged to obtain expert opinion. Even where the capability exists, independent reviews from time to time should be encouraged. We cannot emphasise the importance of ensuring data security is adequate. A single failure here will undermine the success of the whole initiative.

It would be worthwhile pilot testing a few cases prior to the legislation being finalised. If it hasn't already happened the Privacy Commissioner should be consulted on this aspect of the legislation.

Questions about data safeguards:

Is the Five-Safes framework the appropriate mechanism to ensure data is safeguarded?

Are there any additional safeguards that should be applied?

Are there any instances when the Five-Safes could not be applied?

Is the Five-Safes appropriate when data is shared and used for the specific purposes in the purpose test above?

How should the responsibility for managing risks be shared in the framework?

How would you envisage Five-Safes principles be applied over the life-cycle of data to ensure data safeguards are continually met?

Under what circumstances should trusted users be able to access sensitive data?

Response:

ASSA supports the use of the Five-Safes framework. However, we re-iterate the important role to be played by safe data. This is cost-effective for both the data custodian and the researcher assuming researcher needs can be met by the safe data.

ASSA suggests that an appropriate regulatory framework, drawing on the models indicated by the regulatory pyramid literature, be established to support and enhance the Five-Safes Framework and the Purpose Test. Recognising the proposed legislation is principles-based, the regulatory

framework could be set out at a principles level. One example is the important role that education will play in ensuring that 'trusted users' understand the implications of specific actions in the context of their work, and in ensuring that trust in and compliance with the Five-Safes Framework and the Purpose Test will build over time.

As with any regulatory pyramid mode, responses should escalate with the severity of non-compliance, and in cases of repeated non-compliance. In very extreme case, criminal sanctions may be appropriate. We would suggest this possibility be included in the penalties. It would be a deterrent and if there was an extreme breach, the public would expect criminal sanctions.

ASSA would like to see some further detail on the concept of a 'trusted user', including whether it will apply to individuals or teams or organisational entities, and if teams or organisational entities then how will responsibility for safe behaviour of the team/organisation be assigned. To be a trusted user it should be mandatory for them to have been through a training program (potentially on-line and self-directed) to understand the importance of maintaining confidentiality and security. The ABS experience is that the main risk is that users will share data with other non-approved users or not store data in a sufficiently secure way

We would also suggest that recipients of shared data be required to sign a legally enforceable undertaking. This could be part of the planned templates for data sharing agreements.

It would be useful to consider the implications of multi-party agreements, as these may have different requirements to bi-lateral agreements

Questions about public sector data sharing arrangements:

Would this arrangement overcome existing barriers to data sharing and release?

Would streamlined and template agreements improve the process?

Do you agree that data sharing agreements should be made public by default?

What level of detail should be published?

What else should a data sharing agreement contain?

What other transparency mechanisms could be mandated?

Response:

On the first question, there will be strong cultural barriers to the sharing the data. This is most often middle managers who are concerned about (1) the quality of the data, or (2) exposing the Minister if the data becomes publicly available. It will require special program to overcome this resistance.

ASSA strongly supports the use of streamlined and template agreements. It may be useful to devise several templates relating to factors such as (for example) level of risk and number of parties to the agreement.

ASSA recommends that there should be a clearly set out dispute resolution process for each stage of the public data sharing arrangements. This could include, for example, decisions regarding whether or not an application meets the purpose test, right through to whether the requirements set out in the data sharing agreement were met.

As rightly proposed, the heads of agencies will have the final decision on whether to provide access or not. However, they will need advice to assist them with this decision making. For some agencies, the necessary skills will be available but it would still be worthwhile to ensure they have access to consistent advice by setting up a Review Panel or the like. For some agencies, external experts may be necessary especially in areas like information security.

5. Roles and Responsibilities in the System

Questions:

How long should accreditation as an ADA or Trusted user last?

What could the criteria for accreditation be?

Should there be review rights for accreditation?

Should fees be payable to become accredited?

Is the Australian Government Charging Framework fit for purpose in this context?

Response:

ASSA strongly recommends that accreditation as either an ADA or Trusted User be time limited. The chosen time periods should be implemented based on some comparable processes, and could sensibly include both a brief monitoring type report (activity, significant outputs, changes to the data safeguard arrangements, unforeseen adverse events) on an annual basis combined with a less frequent re-accreditation process. There may be benefit in considering a categorisation of low and high risk users (based on types of data accessed) in order to streamline accreditation processes, in the same way that Ethics Committees frequently implement low and high risk approval tracks.

The question of fees may raise access and equity issues, particularly in relation to trusted users. A major commercial entity may easily meet fees that a small not-for-profit would find prohibitive. This could create an unintentional commercial advantage to larger entities who are in a position to maintain accreditation over time, and charge less well-resourced entities for products or services on a one-off basis. The same argument could be applied to large research intensive universities in comparison to smaller regional ones.

6. National Data Commissioner

This is strongly supported by ASSA. We also urge that the Commission should work closely with the Privacy Commission. It is important to avoid conflicting views.

We support the establishment of an Advisory Board but would argue that it should be skills based in addition to being representative of different institutional sectors. There should be at least one representative with skills in each of statistics, privacy, information management, and information security. In fact, it may be worth setting up an Interim Advisory Board to assist the Department address the many questions that will arise during the drafting of legislation.

ASSA is available at any time to further discuss this submission:

Academy of the Social Sciences in Australia
26 Balmain Crescent, Acton ACT 2601
GPO Box 1956, Canberra ACT 2601
P: +61 2 6249 1788
Dylan.Clements@assa.edu.au