UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Professor Kathryn McGrath
Deputy Vice-Chancellor (Research)
15 Broadway, Ultimo NSW 2007

T: +61 2 9514 1257
Kathryn.McGrath@uts.edu.au

PO Box 123
Broadway
NSW 2007 Australia
www.uts.edu.au

UTS CRICOS PROVIDER CODE 00099F

1 August 2018

Department of the Prime Minister and Cabinet
PO Box 6500
Canberra ACT 2600

RE: Australian Government Data Sharing and Release Legislation: Issues paper for Consultation

TO WHOM IT MAY CONCERN,

The University of Technology Sydney (UTS) is pleased to provide comments on the new Australian Government Data Sharing and Release Legislation.

Unfortunately, we have not had the opportunity to comprehensively respond to each of the questions in the issues paper, as we received late notification of the consultation period. We have provided initial comments in our areas of expertise and would value further engagement.

We understand that Universities Australia has requested further consultation on the legislation and we support their leading engagement with the higher education sector.

Thank you for the opportunity to comment.

Yours sincerely,

Professor Kate McGrath
Deputy Vice-Chancellor and Vice-President (Research)

Overall comments:

We have provided feedback below. The most important points to which we would like to draw attention are:

1. In order to maximise discoverability, the National Data Commissioner, or another agency should be responsible for maintaining a registry of data that (a) exists within agencies or (b) has been shared within the government or (c) is available for sharing for projects that meet the five safes test or (d) has been made available publicly. (See our response to question 11)
2. The special status of universities as trusted organisations with a mission aligned with public good should allow for streamlined DS&R, via commonwealth funded projects (see 8-10), and the definition of research should be broadened to encourage strategic basic research (See 3).

There are two important omissions from the issues paper:

1. We note that the issues paper does not include a definition of data. We strongly urge that data is defined as broadly as possible so that it includes not only tabular (database-type) data but all types of data including still and moving images, text and data-streams such as environmental sensor networks.
2. A data-licensing framework is also essential: there needs to be a shared language for expressing what can be done with data and how long it needs to be kept that can be used for cases from unconstrained open-release through to data that is only available in a secure facility.

1. Are these the correct factors to take into account and to guide the legislative development?

UTS commends measures to "improve the sharing, use and reuse of public sector data while maintaining the strong security and privacy protections the community expects." Recent reactions to the introduction of "Opt out" requirements to MyHealthRecord illustrate that misreading community expectations can result in significant missed opportunities for research and innovation of public benefit. The emphasis on "building trust in use of public data", enhancing integrity and transparency and "safeguarding data consistently and appropriately" are absolutely crucial at this point in time and essential prerequisites to the agenda.

As an academic research institution with particular strengths in translational research, UTS would appreciate opportunities for its researchers to access public data under provisions for "research and development with clear and direct public benefits".

Further clarification of "research and development" and "clear and direct public benefits" would be helpful, particularly as the broader community currently seem to have little tolerance for ambiguity about who can access their data and for what purposes.

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

Unfortunately, recent events around MyHealthRecord have amplified a public trust-deficit in government data sharing, even where it can be fairly readily demonstrated to be in the public interest (as with health and medical research). In this context, the implementation of DS&R and other elements of the Data Governance reform agenda must be carefully managed. Benefits will need to be well communicated and questions addressed promptly and clearly.

3. Should the scope be broader or narrower?

The scope should be clarified to unambiguously include data sharing with universities and research institutions as these institutions have demonstrated experience dealing with sensitive data to conduct research in the public interest.

For example, UTS researchers have previously demonstrated capacity to pass audits on managing sensitive data from NSW government agencies.

The issues paper list the following benefits:

- economic growth from innovative data use, and
- cross-sectoral research solutions to current and emerging social, environmental and economic issues.

In this context the wording “research and development with clear and direct public benefits” it problematic at it may be seen to exclude two categories of research under the *1297.0 - Australian and New Zealand Standard Research Classification (ANZSRC), 2008*: (a) Pure basic research and (b) Strategic basic research. Strategic basic research is fundamental to furthering the public good. This is the ANZSRC definition of strategic basic research (with added *emphasis*):

> **Strategic basic research** is experimental and theoretical work undertaken to acquire new knowledge directed into specified broad areas in the expectation of practical discoveries. *It provides the broad base of knowledge necessary for the solution of recognised practical problems*.

There is a high risk that potential benefits that would arise from strategic basic research will not be realised under the proposed DS&R scope.

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

If “data categories” refers to existing information security classifications, this might be sensible, however it may also provide a perverse incentive to over-classify data.

There may be some instances in which a combination of both may be required.

8. Do you agree with the stated purposes for sharing data? and

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

10. What further detail could be included in the purpose test?

The stated purposes for data sharing under point 4 remain somewhat unclear, particularly with respect to “clear and direct public benefits”. It is noteworthy that the term “public interest” which has some status in law, is not used. It may be preferable to utilise terminology where there are existing guidelines for interpretation.

Clarification of the term “research and development” may also be beneficial, given current public concern about sensitive data being shared for profit rather than public interest. A relevant case is recent public reaction to for-profit data sharing by HealthEngine, which, although it *might* yield indirect economic benefits, was not seen as sufficiently in the public interest to warrant sensitive data sharing. Risks like this can be mitigated via a research-first approach to development and innovation which universities, as research institutions, are uniquely well-placed to deliver. Universities have a social license to produce public goods through knowledge creation and translation and social and economic capacity building and university research which is undertaken under the governance should be included, unambiguously, in the purpose test.

IMPORTANT: Research with a research-council grant or other commonwealth funding, including but not limited to ARC and NHMRC funded research projects should automatically be deemed to meet any purpose test under the DS&R, assuming the five-safes protections are met.

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

Yes, a discovery service is needed so that potential users can learn what data exists within the departments and agencies covered by the DS&R bill - even if only metadata describing the data can be shared publicly. Ideally metadata could include information about the structure of the data, indicative content and who to contact, or what process to follow to apply for access to the data.

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

Where data is not sensitive and does not require de-identification, it should be shared publicly by default, but note that there is no mechanism under the proposed DS&R scope to ensure that this happens.

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

It is unclear here whether "compliance requirements" refers to compliance with this legislation, which would be appropriate, or compliance on a broader scale with government legislation and policy. The latter may be appropriate in some circumstances, but has potential to become a sensitive issue if the public perceives that it is being deployed to benefit party political interests rather than public interests.

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

Yes, the Five-Safes framework is an appropriate mechanism with sufficient flexibility, and the ABS application of it is a good example.

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

Yes

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

While the arrangement may overcome barriers by mitigating risks to agencies who share and release data, questions remain about the extent to which the arrangement creates "incentives for self-management of risks and voluntary improvement of data management practices within public sector agencies." The arrangement creates a new and (in some cases, necessarily) complex process which, if not adequately resourced, will itself be an impediment to DS&R.

We note that "The DS&R Bill will not by default compel all data to be shared but rather will support data custodians to make informed decisions and manage risk consistently to enable appropriate sharing and release." We accept that compelling data sharing cannot be the default position for sensitive data, but suggest that it should be for non-sensitive data (with the proposed exclusions). It appears from this statement "The sharing of data is at the discretion of the data custodians.", that the DS&R bill **will** not compel data sharing in any circumstance - even if due process is followed and a recommendation is made by the NDC to share the data.

Long experience in trying to develop a culture of data sharing among researchers suggests that an arrangement with neither carrots nor sticks will not significantly improve data sharing.

22. Would streamlined and template agreements improve the process?

Yes. Streamlining all aspects of the process will be absolutely crucial. This raises the question of whether there could be a mechanism for establishing precedents for overriding secrecy or confidentiality under DS&R.

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

24. What level of detail should be published?

Yes, but not to the extent that they might indicate the location of a “honeypot” to would-be cyberattackers.

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

In both cases accreditation should continue as long as audits are passed. For Trusted users, a non-compliance event of any seriousness should revoke accreditation.

28. What could the criteria for accreditation be?

For ADAs: it is proposed that “criteria and process for accreditation would be developed by the NDC and could build on requirements for accrediting Integrating Authorities.” We note that currently there are only three integrating authorities and this is probably a potential choke point in the new DS&R ecosystem. We recommend that more potential ADAs are identified and funded to upskill if necessary.

It should be noted that the Federal government (for example, under NCRIS and NISA via ARDC) have already funded some subject-specific data sharing infrastructure (for example, Australian Data Archive, Terrestrial Ecosystems Research Network, etc). Since there are significant discovery, re-usability and interoperability advantages to sharing data via these repositories, it could be beneficial to consider whether they should be considered for ADA status.

For trusted users: Only two of the five safes (people and setting) can be assessed independent of the specific case (project, data, outputs are all request-dependent) and neither ‘Safe people’ nor ‘Safe setting’ are adequate without the other. Therefore both together should be demonstrated to determine Trusted User status.

29. Should there be review rights for accreditation?

Yes

30. Should fees be payable to become accredited?

This would potentially disadvantage the research sector which already operates in straitened circumstances. Public funding agencies would need to decide whether these costs can be included in grant costings. Ideally accreditation fees would recover costs of accreditation, rather than become a means of recouping other costs of the DS&R ecosystem.

32 - 43. Questions relating to the National Data Commissioner and processes for non-compliance

The information provided in the issues paper relating to questions 32 to 43 indicates that significant consideration into the role of the National Data Commissioner and processes for handling non-compliance is yet to take place. UTS welcomes further opportunities to engage with these developments.

As stated previously, universities are trusted and practiced organisations involved in the collection, management and analysis (research) of data. Universities and other research organisations have comprehensive processes and governance arrangements around the handling of data for research purposes, aligned with the *Australian Code for the Responsible Conduct of Research*. It will be important for any approaches developed in this space to be mindful of existing mechanisms for investigation and management of non-compliance, to avoid separate and conflicting processes, and to ensure that universities are able to meet their existing responsibilities under the Code.