26 July 2016

To whom it may concern

    RE: Productivity Commission: Data Availability and Use

Please find a submission to the Productivity Commission Issues Paper on Data Availability and Use.

This submission is being made on behalf of the Atlas of Living Australia.

Thanks you very much for your consideration of this submission.

Sincerely.

Dr John La Salle
Director Atlas of Living Australia

Atlas of Living Australia

Submission to Productivity Commission: Data Availability and Use

Key Messages

- The Atlas of Living Australia strongly supports an Open Data policy (at least for biodiversity and environmental data) – and continues to demonstrate the success of making biodiversity and environmental data freely and openly available.

- The ALA provides solid examples of the benefits and impact of making data open and available. The value of data improves when it can be integrated with other data within an environment that supports data discovery, visualisation and analysis. The success and achievements of the ALA have help to drive a culture shift in terms of data sharing.

- The ALA supports an Open Data Policy and feels that any data generated with public funding should be made open and freely available. Having said that, the ALA recognises that there are reasons why a sensitive data service is necessary to protect certain types of data (personal information, data covered under legislation, data around rare or threatened species, biosecurity information that could affect market access). While there should be a quarantine period on data that is protected to provide researchers with time to publish results from that data, there should be strict policies on the length of that period.

- The ALA encourages the custodians of private sector data to share that data through demonstrating the direct benefits to the private sector of data sharing (e.g. improved environmental impact reporting).

- Open Data is only one component of the culture shift that needs to be achieved as we embrace the Digital Transformation. We need to promote open code, open infrastructure, open services and a culture that embraces collaboration rather than the creation of data/infrastructure silos.

- While we support the issues paper call for standardising, sharing and release of data (where for the most part there are existing protocols already in place), we would like to see additional emphasis on the creation of data. This should include new methods of data capture, "cradle to grave" data management policies, and rapid digitisation programs to bring physical information into the digital world.

Atlas of Living Australia

Submission to Productivity Commission: Data Availability and Use

Specific responses

Data Availability and Use – Productivity Commission Contribution

1. **Benefits and costs of options for increasing availability of public sector data to other public sector agencies (including between the different levels of government), the private sector, research sector, academics and the community. Where there are clear benefits, recommend ways to increase and improve data linking and availability.**

    a) **identify the characteristics and provide examples of public sector datasets that would provide high value to the public sector, research sector, academics and the community to assist public sector agencies to identify their most valuable data; and**

The ALA combines mainly public sector data sets (over 1000 data sets), with notable contributors being Commonwealth Government Departments, all State and Territory Governments, Australian biological collections (museums and herbaria), NGOs, NRM and CMA groups and Citizen Science initiatives.

The Atlas of Living Australia has achieved a great deal in establishing the benefits of sharing public data sets. These include:

- The ALA currently holds over 60 million records, and to date over 7.5 billion records have been downloaded by a wide user community to support research, education, conservation management, biosecurity, environmental assessment, collection management and citizen science.
- Over 800 academic articles have referenced the ALA since 2007 (Source: Google Scholar)
- ALA represents the largest partnership for biodiversity in Australia
- Digivol – an online platform enabling volunteers to digitise museum collection specimens – has reached over 350,000 digitisation events
- There are over 1 million images and 4500 sound files on the Atlas
- In 2015 the Atlas of Living Australia had over 625,000 users, over 5 million page views, and an average 3500 users per day (Source: Google Analytics)

    b) **examine legislation or other impediments that may unnecessarily restrict the availability and linking of data, including where the costs are substantial, and consider options to reduce or remove those impediments.**

Nil response

2. **Benefits and costs of options for increasing availability of private sector data for other private sector firms, the public sector, the research sector, academics and the community. Where there are clear benefits, consider ways to increase and improve availability.**

   a) **identify the characteristics and provide examples of private sector datasets that would provide high value to the private sector, public sector, the research sector, academics and the community in developing or providing products and services, undertaking research and developing policy;**

While mainly working with public sector data, our largest single data contributor is BirdLife Australia (12.5 million records). Whether an NGO is public or private sector is perhaps debatable, but this is an example of a non-publicly funded organisation that saw the value of contributing its data to the ALA.

   b) **identify the concerns of private sector data owners and provide recommendations on principles or protocols to manage these concerns;**

Nil response

   c) **examine legislation or other impediments that unnecessarily restrict the availability of data, including where the costs are substantial, and consider options to reduce or remove those impediments; and**

Nil response

   d) **provide an update on existing data sharing initiatives in Australia, including the uptake of the credit reporting framework. Consider recommendations for improving participation in such initiatives.**

Nil response

3. **Identify options to improve individuals' access to public and private sector data about themselves and examine the benefits and costs of those options.**

The ALA does not hold these types of data.

4. **Options for, and benefits and costs of, standardising the collection, sharing and release of public and private sector data.**

For environmental and biodiversity data, we have several options for data sharing already in place and working. These could be further developed by expanding the existing collaboration between the NCRIS funded environmental facilities (e.g. ALA, IMOS, TERN). Data standards are already in place (and widely used) for much of this information.

Some government departments still maintain some of their data separately from the NCRIS environmental initiatives. While holding and managing the data separately may be sensible, the sharing of the data would be of mutual benefit.

While not specifically called for, the ALA would add the following comments to this section:

- Open Data is only one component of the culture shift that needs to be achieved as we embrace the Digital Transformation. We need to promote open code, open infrastructure, open services and a culture that embraces collaboration rather than the creation of data/infrastructure silos.
- Examples of benefits here include:
    - The core ALA infrastructure now supports over 70 different hubs and portals delivering to a wide variety of outcome areas (biodiversity, biosecurity, citizen science). This is a key example of the benefit of open infrastructure – which is as important (if not more important) a story than open data.
    - ALA infrastructure is now being used by 7 different countries around the world to support their own national biodiversity portals – with several more actively implementing the ALA tools.
    - ALA web services (all completely open) are being used by various groups to support their own initiatives (e.g. NECTAR labs such as the Biodiversity and Climate Change Virtual Laboratory; QuestaGame – a game app to support biodiversity learning).
- While we support the issues paper call for standardising, sharing and release of data (where for the most part there are existing protocols already in place), we would like to see additional emphasis on the creation of data. This should include new methods of data capture, "cradle to grave" data management policies, and rapid digitisation programs to bring physical information into the digital world.
    e-Infrastructure such as the ALA is only as good as its content – so initiatives that increase the flow and quality of data are always welcomed.

5. **Ways to enhance and maintain individuals' and businesses' confidence and trust in the way data are used. Having regard to current legislation and practice, advise on the need for further protocols to facilitate disclosure and use of data about individuals and businesses while protecting privacy and commercial interests and, if recommended, advise on what these should be.**

The ALA does not work in this area.