Australian Government
Department of Agriculture
and Water Resources

SECRETARY

Ref: EC16-000479

Chairman Harris AO and
Presiding Commissioner Cilento
Data Availability and Use Inquiry
Productivity Commission
data.access@pc.gov.au

Dear Commissioners

Thank you for the opportunity to respond to the Productivity Commission's issues paper on *Data Availability and Use*.

The Department of Agriculture and Water Resources collects and manages a substantial volume of data. As a policy developer, economic and scientific research adviser, regulator and programme and service provider, the department is heavily reliant on its data assets. The objective of all our activities is to build more profitable, more resilient and more sustainable agriculture, fisheries and forestry sectors, and to support the sustainable and productive management and use of rivers and water resources.

There is an increasing awareness in the rural sector, including among researchers and industry, of the value of public data and an expectation that data will be more freely available. The department's Portfolio Budget Statement 2016-17 identifies that managing public data for the benefit of the Australian people, in line with the government's public data policy, is a significant internal reform.

To progress this reform, I have established a departmental taskforce to establish policy and processes to improve the way the department manages, shares and publishes data resources. The taskforce is working with portfolio agencies and statutory Research and Development Corporations to achieve a consistent approach where possible across the portfolio.

The taskforce, expected to operate until June 2017, will progress various steams of work. These include the development of a departmental catalogue of all departmental data assets, developing and implementing a comprehensive information and data management framework, liaising with stakeholders to prioritise data for sharing and publication and developing a future state vision and roadmap for information and data management. Together, these initiatives will enable the department to implement and maintain policy, processes and infrastructure that will support the sharing and publication of data.

T +61 2 6272 3933
F +61 2 6272 5161
18 Marcus Clarke Street
Canberra City ACT 2601
GPO Box 858
Canberra ACT 2601
agriculture.gov.au
ABN 24 113 085 695

Achieving these objectives is not without challenge. Portfolio legislative provisions that constrain the sharing of information and data, the capability and capacity of our information technology infrastructure, the expertise of staff and the cost of implementing and maintaining this reform are just some of the challenges that will need to be overcome. These and many other factors will influence the extent and timeliness of sharing and publication of our data.

I welcome the inquiry and am confident that the findings will provide guidance to further the department's efforts to embed the public data policy as part of its day-to-day operations. Attachment A to this letter provides a more detailed response to some of the questions posed in the issues paper.

Yours sincerely

Daryl Quinlivan

28 July 2016

Data Availability and Use - Productivity Commission Issues Paper

HIGH VALUE PUBLIC SECTOR DATA

1. *What public sector datasets should be considered high-value data to the: business sector; research sector; academics; or the broader community?*

While some datasets will inherently be more valuable than others because of the information they contain or their application, the value of a public dataset outside government will ultimately be determined by the user and their ability to find application for that data on its own or in combination with private sector data. The government has a role to play in facilitating data discovery and making data accessible where appropriate.

The Department of Agriculture and Water Resources (the department) is keen to develop a common understanding with the business and research sectors on what datasets they consider to be of high value, with the aim of prioritising these for sharing and publication if possible. The department is developing a catalogue of the data assets it holds to better understand the scope of our collection. This information will form the basis of future consultations with our stakeholders. We acknowledge that the way we use our data can have limitations and that others may find a higher value use for our datasets than we realise.

2. *What characteristics define high-value datasets?*

The Australian Government has defined high value datasets as those that are valuable to business/industry, help the public make informed decisions or improve user experiences, or assist government in making evidence-based policy decisions[1].

i. Valuable to business/industry:

The primary industries sectors have a strong history of innovation and capacity to adopt new technologies. The government recognises that investment in research and development and innovation is vital for ongoing growth and improvement in the productivity, profitability, competitiveness and sustainability of Australia's agriculture, fisheries, forestry and food industries.

From the perspective of industry and business, the value of a dataset primarily depends on its accessibility, readability, quality and re-use potential, and in particular, its commercialisation potential. Improved access to public data assets that are re-useable could support the development of new products and/or services that have direct or indirect economic impact and benefits and encourage further scientific enquiry and debate. For the primary industry sectors, datasets that contain productivity information and geo-spatially enabled data are likely to be of high value.

Within existing data sets, considerable value can be added though data-linkage work. We note that relatively little data-linkage has been done using a geo-spatial spine.

---

[1] 11 March 2015. Malcolm Turnbull PM (previous Minister for Communications). Call for input into open data study.

ii. Facilitates informed decisions or improved user experiences for the public:

A dataset would be considered of high-value when it contributes to transparency and openness of government to improve experiences for the public. Industry and business rely on critical datasets to locate markets/customers, organise supply chains and analyse performance. Providing access to data, if available, that would assist with these activities could lead to increased economic benefits for these businesses and therefore benefit the economy as a whole.

If public data is to be reusable, then it needs to be machine-readable, of 'known' quality, consistently defined, and meets recognised standards. This would provide greater confidence to the user about the quality of the data. Data quality management is a key priority for the department as it develops a new information and data management framework. The department acknowledges that the availability of public data, underpinned by common standards, has the potential to eliminate duplication of effort; thus reducing costs and increasing interoperability.

iii. Assists government in making evidence-based policy decisions:

Data is integral to the work of the department, as a policy developer, economic and scientific research adviser, regulator and programme and service provider. While the department has not formalised a method to value the data assets it holds, characteristics such as alignment with departmental objectives, currency, frequency of use, ease of access and data quality contribute to the overall perception of data value.

Extending the availability of non-sensitive data would support greater public participation in policy development and decision-making. The key to evidence based decision making is using knowledge produced through multiple sets of data and analysis to inform and influence policy rather than determine it[2]. Allowing industry and business greater access to our datasets would provide them with greater ability to influence policy decisions.

COLLECTION AND RELEASE OF PUBLIC SECTOR DATA

3. *What are the main factors currently stopping government agencies from making their data available?*

The public data agenda presents a considerable challenge for the department, requiring consideration of a range of legal, information and communication technology and resource implications.

Legislative barriers

The Minister for Agriculture and Water Resources is responsible for administering in excess of 95 Acts[3] and many pieces of subordinate legislation. Many of these contain controls that affect the way in which information (including data) can be gathered, used,

---

[2]OECD Public Governance Reviews Finland: Working Together to Sustain Success 2010

[3] https://www.dpmc.gov.au/pmc/parliamentary-information

shared and published. While many of these controls are necessary and appropriate, legislation in general has not been drafted with the view to share or make data public.

In addition to our portfolio legislation, there are a number of laws that affect how Australian Government agencies manage information. These include the *Freedom of Information Act 1982*, the *Privacy Act 1988* and the *Public Governance, Performance and Accountability Act 2013*. To realise the full value of data, unnecessary legislative impediments to sharing and publication of data need to be managed at a whole of government level and a consistent approach to the drafting of all Commonwealth legislation adopted.

The department has coercive powers under legislation to collect information and material, including notices to produce documents or to answer questions. The High Court decision in *Johns v Australian Securities Commission* sets out when information collected under such statutory powers can be disclosed and how it can be used. The principle in *Johns* is that information collected by an agency under a coercive power may usually only be used or disclosed for the purpose for which that power was conferred, subject to any express statutory power permitting otherwise. The information obtained must be treated as confidential, whether or not it would otherwise be confidential. The person to whom to the duty of confidentiality is owed must be given procedural fairness before the information is used.

Information and Communication Technology barriers

The Australian Government public data policy[4] requires government entities to ensure all new systems support discoverability, interoperability, data and information accessibility and cost effective access to facilitate access to data.

Much of the department's ICT infrastructure was built at a time when public data sharing was not a primary consideration and to a large extent, these legacy systems have limited capability to support the policy as currently defined. While there has been significant investment and improvement in information and data management within the department in recent years, including the development of a departmental enterprise architecture and implementation of an enterprise data warehouse (and associated analytics capability), continued investment in ICT capability will be required to support greater data availability and use. This would include upgrades or replacement of existing ICT infrastructure as well as expanding capability in a number of areas, such as the development of application programming interfaces.

Capability and capacity constraints

The growth in volume and complexity of data and advances in the way in which data can be used and analysed requires more specialised capability and capacity of staff to manage data and capitalise on its value. The public data policy, particularly the public release of data and secure sharing outside government, will add a further layer of complexity to ensure that security, privacy and commercial confidentiality is maintained. This work will require additional resources with specialised data

[4] https://www.dpmc.gov.au/sites/default/files/publications/aust_govt_public_data_policy_statement_1.pdf

management skills and increased awareness and development of relevant skills for all who manage and access data.

4. ***How could governments use their own data collections more efficiently and effectively?***

Better integration of the department's ICT systems, access to the data collected by these systems, and the use of new and emerging technologies is crucial for efficient and effective data utilisation. For example, the effective management of biosecurity risks relies upon strong biosecurity measures and controls that serve as a national insurance policy to protect the economy, agricultural production, the environment and human health from pests and diseases.

Better access to, and integration of, a range of the department's existing data holdings such as pest and disease data, trade and transport data, farm production data, habitat and environmental data and social media data would significantly improve the department's management of biosecurity. Analysis drawing from this data would help reduce the risk of pest and disease incursions and improve the effectiveness of mitigation and eradication responses. Linked information could support analytical modelling to determine the highest risk locations and pathways for a potential incursion of a pest or disease based on habitat or farm production data. This would significantly improve the department's capability to predict the economic and environmental impacts of exotic pests and disease incursions and support biosecurity decision making and the allocation of the department's biosecurity management resources.

5. ***Should the collection, sharing and release of public sector data be standardised? What would be the benefits and costs of standardising? What would standards that are 'fit for purpose' look like?***

Yes. The public data agenda is not just about the government putting meaningful data into the public domain, it is also about enabling the public to meaningfully engage with the data and with the government. This will require government conformance with, and promotion of, public data standards (software interoperability, data and document formats), to ensure that the public is able to use the data.

There are many benefits associated with a whole of government approach to data standardisation. It would prevent duplication of effort associated with the definition of these standards and would support consistency, resulting in more readily useable public data.

In the absence of whole of government data standards, the department is developing standards for the management of data across the data lifecycle, aligned wherever possible with applicable Australian and international standards. These standards support the effective management of data during its creation, storage, processing and maintenance, publication and use. Restricting standardisation to these domains means that data is managed in a consistent way, supports data quality and usability but does not impinge on the application of specialised data domains (e.g. standards that relate to spatial data, taxonomies) necessary to ensure data is fit for the purpose it is collected.

6. *What criteria and decision-making tools do government agencies use to decide which public sector data to make publicly available and how much processing to undertake before it is released?*

To date, the overriding considerations for the department when considering the release of data (of its own volition or in response to a request for data) is whether there are any legal impediments to its release. Other risks requiring consideration include the potential for data to be misinterpreted and represented, potential for breach of privacy and or unintended commercial impacts. The resource intensive nature of responding to data requests and undertaking necessary extraction and processing of data, in particular ad-hoc requests from a single user affect the department's capacity to respond in a timely manner.

While these considerations remain more or less unchanged as a result of the public data agenda, the department's decision making and processes for releasing data will be streamlined, supported by a comprehensive information and data management framework and associated governance arrangements. The department is also developing a comprehensive catalogue of its datasets to assist with the management of data assets and support discoverability. This work will support planned discussions with stakeholders to identify data they consider to be 'high value' and assist with prioritising data for publication.

RESOURCE COSTS OF ACCESS

7. *How should the costs associated with making more public sector data widely available be funded?*

Legislative constraints aside, developing ICT capability and capacity to support the public data policy and ensuring data handling and management roles are filled with suitably skilled people could incur additional costs. Over time, some ICT limitations may be overcome as existing systems are replaced with modern ICT infrastructure. Without dedicated additional funding, progress in this area is likely to be slow. Improvements in data management required to meet the objectives of the policy will represent an additional cost to the department in the short term, particularly while processes are established and become business as usual.

Public sector bodies overseas have witnessed a strong growth in demand for information they provide after switching from cost-recovery pricing of data to free or maximum marginal cost priced information. Steenbergen and Carrera assert that cost-recovery of data models do not bring additional revenue to government agencies in the long run while free or marginal cost models are more beneficial from a value creation perspective[5].

The department cost recovers in a number of areas, including biosecurity and export functions, as part of its funding strategy. Implementing a cost recovery framework for public data would add to the administrative complexity of managing data access and

[5] 19 February 2016 van Steenbergen, E. and Carrara, W. on behalf of the European Data Portal *Stop Charging for Open Data*

would potentially be costly to establish and maintain. These costs, on top of the cost of managing data for public release, would inevitably have to be passed on to consumers.

A key element of establishing a cost recovery framework, where full cost recovery is the objective, is understanding the anticipated uptake of a service and correctly attributing costs to users. In circumstances where the number of users is unknown, or cannot be reliably estimated, this attribution of cost is unlikely to be accurately reflected in the fees and charges imposed. Where the data generated is under-used, users may incur high costs. For many, particularly those outside the commercial sector, these costs may prove to be a significant barrier to data access.

8. To what extent are data-related resources in agencies being directed towards dealing with data management and access issues versus data analysis and use?

We recognise that meeting the objectives of the public data policy and capitalising on the value of our data assets (particularly in the field of data analytics) will require significant investments in ICT infrastructure and people. Balancing these demands with necessary business as usual functions limits the extent and speed that accessibility and improved use can be achieved within current resource constraints.

9. Is the availability of skilled labour an issue in areas such as data sciences or other data specific occupations? Is there a role for government in improving the skills base in this area?

Yes. There is a role for government in improving data literacy as this lack of capacity is a significant barrier to the effective use, interpretation and management of data. In the short term, government agencies could focus on re-training people who have some data management skills, and developing procedures for data management. However, to ensure sustainability, specific recruitment targeting data management skills should be undertaken.

PRIVACY PROTECTION

10. What types of data and data applications (public sector and private sector) pose the greatest concerns for privacy protection?

The following data and data applications pose the greatest concerns for privacy protection for the department:

- databases which are accessible by both the public and private sector – for example the Import condition database (**BICON formerly ICON**) which is a database where both department officers and the public access information about Australian import requirements for a range of commodities;
- databases and information held by delivery partners and contractors; and
- sensitive information held about employees and the community with whom the department deals and relies on having relationships of trust and confidence.

The challenge is to de-identify data so that it still has meaningful value. However, for some areas of operations de-identification of information may not be practically achievable or would involve significant work to ensure it is effective.

11. How can individuals' and businesses' confidence and trust in the way data is used be maintained and enhanced?

The department is very concerned to ensure that individuals and businesses have trust and confidence in the department's data handling. It maintains a current privacy policy, reviews its personal information practices and provides privacy notices particular to the information being collected.

The department also has:

- appointed privacy officers to respond to privacy complaints;
- a legal branch that provides advice, guidance and training on information handling; and
- need-to-know database access permissions to conform to privacy, secrecy and confidentiality requirements.

The department takes steps to protect the security of the personal information it holds from internal and external threats by:

- regularly assessing the risk of misuse, interference and loss, and from unauthorised access, modification or disclosure of that information;
- conducting internal and external audits of our information management and security practices; and
- addressing those risks through:
    - the department's Protective Security Policy and Plan, which outlines the governance arrangements around security matters, including information security, physical security and security incident management;
    - the department's Recordkeeping Policy, which outlines how the department ensures information (including personal information) is managed in an efficient, consistent and accountable manner;
    - the department's IT Security Policy and Handbook, which outlines how the department holds and secures electronic information (including personal information); and
    - other Commonwealth standards including the Digital Transition Policy, and the Defence Signals Directorate Guidelines, including the Information Security Manual.

12. What weight should be given to privacy protection relative to the benefits of greater data availability and use, particularly given the rate of change in the capabilities of technology?

Note comments above about challenges of de-identification of data. However, the department notes that it is important to be able to appropriately share information with other agencies with related functions and roles to the department in the agriculture space. To that end, the department seeks to ensure that it has appropriate mechanisms to share information within privacy and other legislative constraints.

13. How could coordination across the different jurisdictions in regard to privacy protection and legislation be improved?

Privacy requirements are not generally an impediment to engagement with other jurisdictions. However, effective disclosure provisions to permit appropriate information sharing are an important consideration when legislation is being developed or reviewed, particularly where secrecy provisions are involved.

14. How effective are existing approaches to confidentialisation and data security in facilitating data sharing while protecting privacy?

Refer to above comments about data security practices and de-identification of data.

The department also maintains, or has access to, a number of specialised database systems with varying levels of restricted access depending on the sensitivity of the data.

15. What competing interests (such as the public interest) or practical requirements would indicate that the ability to request deletion should not apply?

To the extent the information is contained in a Commonwealth record, the *Archives Act 1983* will also apply and imposes additional requirements relating to destruction of that information.

There are also practical difficulties in responding to deletion requests where it is difficult to identify all the instances of the data requested to be deleted.

DATA SECURITY

16. How do data security measures interact with the Privacy Act?

The department recognises its requirements under Australian Privacy Principle 11 to 'take such steps as are reasonable in the circumstances' to protect personal information. As noted above, the department is very concerned to ensure that individuals and businesses have trust and confidence in the department's data handling and takes steps to ensure risks to data security are appropriately managed.