29 July 2016

Productivity Commission
GPO Box 1428
CANBERRA CITY ACT 2601

Re: Data Availaibity and Use

This submission to the Productivity Commission Data Availability and Use Public Inquiry seeks to broadly address the Terms of Reference within the context of the role of public sector in making data more discoverable and accessible.

The Centre for eResearch and Digital Innovation (CeRDI) at Federation University Australia (FedUni) is a global leader in data interoperability and has developed many award-winning web-based information and knowledge portals which provide free public access to data sets that are often hidden from view[1]. The portals meet end-user and research partner needs across areas which include Agriculture, Natural Resource Management, Heritage and Culture, Hazards Planning and Resilience, Health and Wellbeing and Regional Development.

Open and interoperable data approaches are central to digital transformation and practice change, to increasing and improving collaborations and knowledge transfer and to enhancing planning and decision making across industry, government and research. Recent increases in the availability of public sector data including census data and other public sector data such as environmental, climate, land use, health and research data have been central to the research and data sharing initiatives we are leading. However, profoundly beneficial outcomes will by generated by continuing to increase public and private sector data availability and use.

CeRDI has been active in supporting its research partners in making use of open data toolkits such as Data.gov.au which provides quality hosting infrastructure, publishing tools and data discovery services for local governments and other public sector organisations. The City for Ballarat has for example published 69 datasets via data.gov.au. These datasets have then been accessed and reused in initiatives such Historic Urban Landscapes and Visualising Ballarat[2] and support Ballarat's continued state and national success in GovHack.

In other areas, much of the public sector data that is accessible online has limitations in that it is not easily discoverable, not spatially searchable (i.e. via a web-GIS interface) and limited by the format in which the data is delivered. For example data is locked into file formats (*.pdf) which are not globally searchable and therefore limited in their usefulness. Nor are the reports and the data they contain visible on geographic maps, which also restricts their discoverability, especially as user may have to know a specific property address, local government area or similar to search and discover relevant information. By contrast, valuable resources such as the following make public sector, industry and community information much more easily discoverable using a combination of spatial mapping, single source publishing and other knowledge management approaches:

---

[1] See CeRDI annual report http://www.cerdi.edu.au/cb_pages/annual_reports.php and FedUni Spatial http://spatial.federation.edu.au/
[2] See www.hulballarat.org.au and www.visualisingballarat.org.au

BALLARAT   PO Box 663 Ballarat VIC 3353 Australia
GIPPSLAND PO Box 3191 Gippsland Mail Centre VIC 3841 Australia
WIMMERA   PO Box 300 Horsham VIC 3402 Australia

EMAIL info@federation.edu.au
CALL 1800 FED UNI | WEB federation.edu.au
ABN 51 818 692 256 | CRICOS Provider Number 00103D

- Visualising Victoria's Groundwater[3]
- Corangamite Soil Health Knowledge Base[4]
- Geelong-Queenscliff Our Coast[5]
- South West Climate Change portal[6]

Portals which federate all publically accessible information around interest areas such groundwater data for Victoria, make legacy data, government datasets, research data and community sourced data and observations visible to the public are effective in meeting the end user needs of farmers, water managers, land contamination consultant and others. They are innovative in terms of offering real-time access to remote authoritative databases by integrating the interoperable web services they each provide. The inclusion of tools for data querying and 3D visualisations are designed to meet end-user needs and educate the broader community about resources that are normally invisible.

CeRDI's unique model of eResearch includes data discovery research, technological research and longitudinal impact research. The impact of the Visualising Victoria's Groundwater web portal was measured using multidisciplinary research that employed survey instruments, expert reference groups, and internet analytics to explore the extent to which the web portal has supported decision making by governments, industry, researchers and the community. The research found that single access, multiple data set web portals enhance capacity by providing timely, informed and accurate responses to answer queries and increase productivity by saving time. Thus, the provision of multiple datasets (including public sector data) from disparate sources within a single portal has changed practices in the Victorian groundwater industry[7].

The public sector clearly has a role as the custodian for government, environmental and social data for Australia, including data curation and management. It is our contention that the public sector also has a role in making this data discoverable to interoperable web portals so that it can be integrated with the plethora of other data and used to enhance decision making across government, industries and communities. Such examples extend beyond the soil, water and natural resource management portals developed by CeRDI, to include others that are used by a range of private industries, public agencies and research institutions. In addition, these data can feed into web-based tools or mobile applications that will dynamically generate decision useful information such as environmental indicators for any place in the landscape, agricultural and land quality insights, societal and health indicators, groundwater quality, air quality, and so on.

While possibly desirable from an agency point of view, it is not essential that the public sector organisations make their data available via spatial mapping portals. We argue that it is more important that data is made discoverable and interoperably available, preferably via web services based on open standards. For most (perhaps all) applications, the key value in public sector data is only apparent once it has been combined with the exponentially growing plethora of data, much of which is collected by sensors (e.g. satellites, in-situ probes, sensors, remote sensing drones, etc.) The value of making public sector data openly available, especially in a way that ensures its currency, authority and integrity, has been clearly demonstrated through CeRDI research.

---

[3] www.vvg.org.au

[4] www.ccmaknowledgebase.vic.gov.au/soilhealth/

[5] www.ourcoast.org.au

[6] www.swclimatechange.com.au

[7] See peer-reviewed scientific paper detailing the results of this research in the Journal of Hydroinformatics, available at: www.iwaponline.com/jh/up/jh2015169.htm

Hence with improved data curation, management and provision of data via interoperable web services for consumption in external web portals, the public sector could:

- Increase availability and use of data and promote a diverse range of benefits across industries, government, research and communities;
- Improve regulatory efficiency and minimise regulatory burden by improving the discoverability and interoperability of data so that it becomes decision-useful and hence empowers industries and individuals with the capability to make timely, informed and accurate responses while increasing productivity by saving time;
- Support new approaches for the public sector to carry out its role in relation to areas such as economic, public health and environmental issues by responding with evidence based information in areas of community concerns;
- Better meet the Australian community's and industry's expectations of the public sector and improve accountability of governments;
- Foster innovation and new market opportunities by making data available for new products and process.

Change management, capability and capacity building investment will be required within and external to public sector. Staged and targeted approaches around specific industry, societal or environmental use cases and which focus on the delivery of data analysis and decision support will deliver earlier benefits that approaches which just focus on making public sector data more widely available.

'Business as usual practice' across the public sector must in future result in making public sector data widely available. While data discovery and publishing costs will likely be borne by the public sector the data analysis and innovation investment is likely to be driven by research which explores discoveries in the data and/or by the private sector through the development of innovative decision support or other tools and technologies that make use of public sector and other data.

Should further clarification be required, please do not hesitate to contact me.

Yours sincerely,

Associate Professor Helen Thompson
Director Centre for eResearch and Digital Innovation