



Australian Government



AUSTRALIAN INSTITUTE
OF MARINE SCIENCE

TOWNSVILLE

DARWIN

PERTH

5 August, 2016

Online Submission

Mr Peter Harris AO
Chairman

Melinda Cilento
Commissioner

Dear Mr Harris and Ms Cilento

Productivity Commission Issues Paper: Data Availability and Use

Thank you for the opportunity to comment on the Productivity Commission's Issues Paper on Data Availability and Use.

By way of introduction, the Australian Institute of Marine Science (AIMS) is Australia's tropical marine research agency. After decades of oceanographic, geochemical, molecular and genetic research in Australia's tropics we now have substantial data holdings that extend throughout Australia's tropical marine jurisdiction. AIMS is a Publicly-Funded Research Agency with statutory independence granted by the AIMS Act (1972). Under the PGPA Act 2013, we are a Corporate Commonwealth Entity that is partially publicly funded with the remainder of our funds coming from co-investment including from various industry sectors. This means that we bridge the private and public sector dichotomy and thus are familiar with the complexities of multi-party data ownership as well as the potential value that well-managed, properly structured and curated data can deliver for data stakeholders. From this background, we offer the following comments and submissions for your consideration.

Data taxonomy

The data taxonomy used in the Issues Paper is useful and clear but we recommend that an additional category be recognised, namely modelled data which is beyond structured data and not simply data visualisation. Many organisations, such as the Bureau of Meteorology, generate substantial data through models. In our case, we generate significant data holdings through oceanographic modelling (which ingests structured data as per the definition of this term in your taxonomy) and these models then forecast conditions at locations which are not monitored and for which primary data is not available. This data is as valuable as data collected directly and in our view the same principles should apply.

Value of Data

In addition to the points raised in the Issues Paper we believe that the most valuable datasets are those that have been developed after consultation between data providers and data users/stakeholders. This ensures that the data that is generated will be immediately relevant and able to be used as soon as practicable after generation. Such close engagement between data user and

provider would be promoted by our submission below for a national approach to marine environmental data.

We also note that the value of a dataset can in certain circumstances meaningfully increase over time. This is particularly the case with environmental data which enables agencies like ourselves to attribute observed environmental changes to both natural long-term processes and anthropogenic activities. Maintaining these long-term time series with in-built controls to ensure their value can not only be maintained but enhanced is critical to our ability to be able to provide the best possible advice to both government and industry and we believe it is essential that longevity be appropriately recognised and preserved.

A National Approach to Marine Environmental Data

AIMS strongly believes that there should be a national approach to marine environmental data. This would realise considerable value for both Australian industry and government. For example, there would be more timely and less costly environmental approvals. There would also be improved transparency for industry development in our marine territory thereby increasing public confidence, and enhancing industries' social licence to operate. Government would also more easily be able to prove it is fulfilling its obligations for conservation and management of the living resources in our EEZ.

The marine community has taken some steps towards a national approach with Geoscience Australia having a well-established system for marine geophysical data, while both the CSIRO and ourselves have well developed systems to catalogue and archive marine environmental data from our ongoing field programs. All of these systems link to the [Australian Ocean Data Network](#) (AODN) that provides a single portal into marine data at a national level. However, these initiatives need increased engagement from the university sector and greater industry participation. This initiative could also drive improvement in the quality of data through increased standardization of sampling protocols and procedures, improved quality control, and greater attention to requirements for spatial and temporal replication. Collectively this will also allow identification of critical gaps in our understanding to inform future marine science investment.

Solving some of the roadblocks to a National Approach to Marine Environmental Data

In our view, the key roadblocks are 1) under-resourcing, 2) skill shortages, 3) security/sensitivity and 4) perceived undermining of business competitiveness.

Under-resourcing - Considerable effort is required to develop data management and delivery tools that ensure the data is not only usable for the initial purpose but reusable for perpetuity. Many organisations recognise this and invest considerable resources in this endeavour but it remains insufficient. In our own case, we have a dedicated team to develop systems to ingest, manage and deliver our data, including to national research data portals. Harmonisation of procedures and adoption of standards reduces the overhead of managing data well, but increased national investment is still warranted. We submit that this could be funded by cost sharing models that include co-investment from data users/stakeholders which commences when data collection starts, not afterwards which is currently more likely if it occurs at all.

Skills - The number of data scientists and analysts needs to grow rapidly. Government can and should redress this shortfall by providing incentives via the education sector for individuals to enter these professions and grow the pipeline of students from which these individuals can be drawn. These skills should extend to include the data delivery skills because without it, the value of the data is often difficult, if not impossible to be derived. As an example, the lack of a commercial off-the-shelf and flexible system suitable for managing and delivering environmental data prompted us to develop our own (open source) system (the [eAtlas](#)). This tool is an efficient and effective way of describing existing intellectual assets and communicating data holdings which has significantly contributed to an increase in data use and innovative integration of datasets.

Security and risk - There are often sensitivities about potential misuse or malicious alteration of data post its download if it is made widely and openly available. Often the level of risk is more perceived and precautionary than real; however this is not to say it is non-existent. Continued education and close engagement between public and private sector organisations with common data interests will build a higher degree of trust and lead to data owners with these concerns lowering their obstacles to data sharing where no feasible justification exists. This would be another benefit of the abovementioned national approach to marine environmental data.

Undermining business competitiveness - While the potential benefits of data sharing and open access to environmental data are acknowledged by industry, the data is often viewed as providing a competitive advantage and may be tradeable. A national approach will provide the required critical mass and scale to manage the volume of potentially confidential data held by industry and manage the access conditions to confidential datasets and the process of releasing data more broadly as concerns about confidentiality wane or provisions expire.

In summary, our submission is that open and easy access to data is critical to Australia and this underpins our approach to our own data and support for the Australian Government Public Data Policy Statement (Turnbull, December 2015) and to managing risks as noted in the Guidance on Data Sharing for Australian Government entities (PMC, March 2016). Our experience is that much of the information from Australia's vast, and largely unexplored, marine territories is often collected by industry and their consultants, is not made public and is as often as not poorly archived and indexed making access difficult, even for the data owners. This is complemented by significant data holdings in the public sector which are captured in organisational data management systems with only a small proportion of this data then held or described in existing national initiatives like the [Australian Ocean Data Network](#) and Research Data Australia. Open and easier access to these spatial and temporal data sets will allow both industry and government to better manage our oceans and coasts and enable sustainable industrial development and maintain ecosystems critical to continued national prosperity. To achieve this as a nation requires ongoing investment in policy harmonisation, standardising collection processes and in growing the infrastructure needed to properly care for the data holdings.

Regards

John Gunn
CEO Australian Institute of Marine Science