Reforming the public sector through evolving our understanding of public sector information to knowledge infrastructure

I believe that reforming and opening up public sector information is one of the highest impact and long lasting options for public sector reform. It is well known that opening up Government data is a critical component to public sector reform. Its benefits are numerous, including its ability to drive positive cultural change in the public sector, reform service delivery, increase Government transparency, make the public service more agile, make evidence based policy a reality, increase Government efficiency, increase Australia's GDP, stimulate innovation and increase citizenry engagement. If the positives are not enough then we are always left with the simple fact that the public paid for Government data and they have a right to it.

Public sector reform comes down to what should be reformed and where funding should be applied. This submission argues that opening up public sector data and seeing knowledge as a form of modern infrastructure should be the top of this list. This thinking is not new and is common to many of the leading public services around the world.

People need power to operate. This power is carried by powerlines and generated by power stations - all of this is infrastructure which the Government helped fund and build. Today people work in a knowledge economy but unfortunately we have old powerlines and no connections to our major power stations.

The concept of knowledge infrastructure helps open up public sector information by treating data as infrastructure. This infrastructure should be robust, built with economies of scale, and contribute to the public good. These concepts are accepted for roads and now we must think about the same concepts for capture and delivery of parliamentary Hansard, elevation data and Government financial records.

The Government spends enormous amounts of resources, both through direct acquisition, and indirectly through grants, on data. However, the Government or the tax payer rarely get a direct return on the investment in that data. One major reason for this is because the Commonwealth Government funds other levels of Government to collect data and then this data is rarely made accessible to the public and generally not even given back to Commonwealth Government.

An example of how Government data should be captured and used is the census. The Government funds the census which is comprehensive, transparent and available to all. The process is regular so that efficiencies can be made between each census by continuously improving collection and analysis techniques. In other cases, such as the Caring for Our Country grants program $2 billion dollars is spent in large part on data capture and delivery. The data is captured in small pockets around the country to different specifications and generally under different licences. Since the data capture is on a small to medium scale economies of scale do not kick in and as a result the data is unnecessarily expensive. At the end of these types of programs a copy of the data is either not provided back to the Commonwealth Government or it is filed within Government in a way which it can never be found by either other agencies or the public.

Concepts of coordinated data capture and knowledge infrastructure are not new, they are common place around the world. For example, of the $12 billion dollars in the Canadian economic stimulus package $2 billion was for Knowledge Infrastructure. An Australian example is the Gershon review which told us centralising similar services can save money while increasing efficiency. The review was endorsed in full by the Government in November 2008.

So what is needed before we can address open public sector information?

On the policy side we must address over-classification, restrictive licensing, restrictive data formats and the lack of policy around open access to public sector information.

On the service delivery side we must address the lack of data maintenance, lack of delivery infrastructure and lack of agreed standards. We must learn to adopt the principle automate or obliterate (this should not only apply to the private sector).

On the technical side we do not need anything but a decision and the resources to act.

The key factor is the current public sector's culture. Government, particularly Commonwealth, views itself as something detached which can settle or start a discussion when it comes to data. We release data with large press conferences when actually collecting and disseminating data is one of Government's basic functions. We must stop thinking of data as something which is released.

Why can I not see the Consumer Price Index live like a stock price? The simple answer is we could. We could allow raw data to be analysed by experts both internal and external to government immediately rather than holding back until we have done analysis. We tend to believe that our value add is required prior to release, but it is questionable whether this is Government's role. We must priorities data accessibility over data enrichment because the public sector norm for only releasing perfect value added data means data rarely gets released.

On implementation we have most of the pieces but no serious leadership or direction on this topic. There needs to be coordination across Government and a decision made about responsibility and structures for implementation. The Gov 2.0 Taskforce, Commonwealth Spatial Data Integration (CSDI) and Data.Australia.gov.au are good steps, but already we have data.nsw.gov.au and other similar sites launching by the day. This means that there is still no single clear point of access to public sector information for members of the public or public servants.

With leadership we can start to link our powerlines to our power stations. In other words we need to link our large data producers to our delivery infrastructure. A key finding from the Gershon review was "There is no whole-of-Government strategic plan for data centres. In the absence of such a plan, the Government will be forced into a series of ad hoc investments which will, in total, cost in the order of $1 billion more than a coordinated approach over a 15-year period." Data centres are the power lines of a knowledge economy and poor power line coordination will cost $1 billion - how much more money will we spend on uncoordinated data capture? The public

sector needs to coordinate data acquisition and delivery through agencies who are experts in acquisition and delivery of data, we cannot simply dump data onto the web. Agencies such as Australian Bureau of Statistics for statistical data, Treasury for economic data and Geoscience Australia for spatial data are already providing this function but not on the scale required.

I make this submission as a young person, a scientist and most of all a public servant involved in service delivery. I also note that not one member of the Advisory Group on Reform of Australian Government Administration comes from any of these backgrounds.

I would like to work in a public service where I as a public servant, or I as a member of the public, can get any information in a manner which is not just easy but enjoyable. I want to be able to use public sector information with confidence and with it updating without having to change anything. I want to make decisions about my work and my life based on factual up to date data and not just anecdotal evidence. I want open public sector information because that is what the world's best public service would have.