

# Open for Business:

## How Open Data Can Help Achieve the G20 Growth Target

---

A Lateral Economics report commissioned by Omidyar Network

**JUNE 2014**



Lateral Economics



OMIDYAR NETWORK  
A WORLD OF POSITIVE RETURNS

## Executive Summary

Open data offers a new start for economic reform and productivity growth. Leveraging the new possibilities of the Internet to make better use of existing resources, it offers us a way of getting past the political economy of winners and losers.

This report illustrates the potential economic value of reinvigorating the open data agenda in Australia and the G20. Our review of the literature, our modeling, and our case study scenarios show how an open data focus could promote the themes of the G20 agenda—trade, finance, fiscal and monetary policy, anti-corruption, employment, energy, and infrastructure—and in so doing achieve more than half of the G20's 2% growth target.

Scaling recent results of the McKinsey Global Institute suggests that implementation of open data policies including in areas corresponding to G20 agenda items could increase G20 output by around USD 13 trillion over the next five years. This would boost cumulative G20 GDP by around 1.1 percentage points of the 2% growth target over five years.

Similar scaling for Australia suggests that more vigorous open data policies could add around AUD 16 billion per annum to the Australian economy. The modeling and case studies in this report validate the magnitude of that figure and suggest that it may underestimate the potential value of open data.

### VALUE CREATION

The more data is opened, the more it can be used, reused, repurposed and built on—in combination with other data—for everyone's benefit. As our economy and society become more knowledge-based, data are core assets, creating value in their own right and driving social and economic innovation, growth and development.

Opening data creates social and economic value in myriad ways by:

- Reducing costs in the provision of existing services both by government and private sector (i.e. doing the same for less cost);

- Enabling new services and improved quality of services; and
- Indirectly contributing to improvements in governance and government services through improved accountability and citizen involvement, both of which engender greater trust in government.

This report explicates these mechanisms.

### POLICY AND THE OPEN DATA AGENDA

Existing open data policy initiatives must be strengthened to prevent government policies dying the death of a thousand bureaucratic cuts in their implementation.

The report also sketches new frontiers that press the open data agenda beyond government and research data. To generate the benefits estimated above, governments would not just more vigorously open data they fund. They would also create conditions in which private organizations opened more of their data to benefit themselves, their associates in their particular information eco-system, and the wider public. Such an agenda has important international dimensions for the G20's consideration, including the potential for cross-national standardisation and harmonisation. Of course all this must be done in a way that protects citizens' privacy.

Opening data should attract overwhelming support, if privacy safeguards are robust. Whereas other reforms often involve painful cuts and winners winning at the cost of losers, there are minimal losers and widespread winners from open data because it simply makes better use of already existing resources.

### OPEN FOR BUSINESS REPORT

This report presents an overview of findings (section 1), reviews the international and local Australian policy context for open government and open science data (section 2), explores the economic value of open data (section 3), outlines a way forward (section 4), and provides case studies on the impacts of open data (annexes).

## About the Authors

**NICHOLAS GRUEN** directed the project. He is CEO of Lateral Economics. After a career in advising government ministers and teaching at ANU, he founded Lateral Economics and Peach Financial in 2000. He chaired the internationally acclaimed Federal Government's Government 2.0 Taskforce and sat on the Cutler Review into Australia's Innovation System. He is chair of Innovation Australia, The Australian Centre for Social Innovation, the Deakin University Arts Participation Incubator and the Open Knowledge Foundation (Australian Chapter). He is also the Patron of the Australian Digital Alliance. The views expressed in this report should be attributed to him and to Lateral Economics but not to any other organization with which he is associated.

**JOHN HOUGHTON** was the lead author. He is a Professorial Fellow at Victoria University's Victoria Institute of Strategic Economic Studies. He has published and spoken widely on information technology, industry, and science and technology policy. He has been a regular consultant to national and international agencies. Previously, he was Research Fellow at the Centre for International Research on Communication and Information Technology, Principal Economist at the Bureau of Industry Economics and the Productivity Commission, and Principal Advisor for State Development Policy to the Victorian Government. In 1998, John received an Australia Day Medal for contributions to policy development.

**RICHARD TOOTH** was a key contributor to a number of the case studies and provided input to other sections of the report. Richard is a Director with Sapere Research Group. He is a consulting economist with expertise in public policy, competition, strategy and regulation. While he works across a range of industries, he has particular experience with regards to water, energy, transport, natural resources, insurance and other financial services. He is a member of the Australian Centre for Financial Studies Insurance Research Reference Group, a research associate with the Centre for Water Economics Environment and Policy at the Crawford School of Economics, Australian National University and an Adjunct Professor with the School of Business, University of Notre Dame Australia.

### **DISCLAIMER**

While every effort has been made to ensure its accuracy, neither Lateral Economics nor Victoria University offer any express or implied warranties or representations as to the accuracy or completeness of the information contained herein. Lateral Economics and Victoria University, their employees and agents accept no liability in negligence for the information (or the use of such information) provided in this report.

## CONTENTS

Executive Summary	ii
About the Authors	iii
Contents	iv
Foreword	vi
Overview	vii
<b>The Value of Open Data</b>	<b>vii</b>
<b>1. Introduction</b>	<b>1</b>
1.1 Objectives and scope	2
1.2 Outline of this report	3
<b>2. Policy context</b>	<b>5</b>
2.1 Economic principles	5
2.2 Policy response	6
<b>3. The economic value of open data</b>	<b>12</b>
3.1 Introductory digression on non-linearity	13
3.2 The evidence to date	17
3.3 Estimates of the value of open data	23
3.4 Summarising these findings	32
3.5 Case studies and examples	33
<b>4. The way forward</b>	<b>36</b>
4.1 Generic opportunities, issues and actions	36
4.2 Beyond Government 2.0 and freedom of information	38
4.3 Open research data	42
4.4 Encouraging open private data	43
4.5 The open data opportunity and the G20	44
<b>5 Conclusion</b>	<b>48</b>
Annexes: Case studies of the impacts of open data	49
Annex 1: G20 Agenda Item – Trade	49
Annex 2: G20 Agenda Item – Finance	51
Annex 3: G20 Agenda Item – Fiscal and Monetary Policy	51
Annex 4: G20 Agenda Item – Anti-corruption	56
Annex 5: G20 Agenda Item – Employment	59
Annex 6: G20 Agenda Item – Energy	63
Annex 7: G20 Agenda Item – Infrastructure	65

## Foreword

Open data offers powerful opportunities for economic and social renewal. And at a time when there is no shortage of hard choices being proposed, making greater use of data should be seen for what it is, an easy choice. It is one of the few resources that is defined by abundance rather than scarcity. Making the best possible use of an existing and increasing resource is not just common sense, it is the closest we can get to generating economic winners without losers. It is not just an economic and social opportunity but a political one.

Thomas Jefferson paid as eloquent a tribute to the potential of open knowledge and open data as anyone today:

*"He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me. That ideas should freely spread from one to another over the globe, for the moral and mutual instruction of man, and improvement of his condition, seems to have been peculiarly and benevolently designed by nature, when she made them, like fire, expansible over all space, without lessening their density in any point, and like the air in which we breathe, move, and have our physical being, incapable of confinement or exclusive appropriation."*

Data on the web offers extraordinary possibilities. The UK was perhaps the first out of the blocks in searching for policy implications with its Power of Information initiatives in the mid 2000s. US President Barack Obama followed suit in 2009, mandating more open data on his first day in office. In Australia, the Government 2.0 Taskforce report was a groundbreaking initiative to rival those of the UK and US; just this year Australian budget data was made available in machine readable form. We are now in a period where people are focussed on finding the value and impact from open data. It is a major focus of the Open Data Institute that Tim Berners-Lee and I created.

As this report highlights, open data is an emerging global resource. The Open Government Partnership and the G8 Open Data Charter both put open data on the map not just as domestic policy but as policy that governments should pursue collectively. International coordination and harmonisation can enhance the value of existing data and open up new global opportunities in ways that unilateral policy making cannot.

The Brisbane meeting of the G20 offers the chance to broaden the open data debate. Most obviously, it broadens the number of countries involved. It can also broaden the agenda itself, organisations of all types, public and private, large and small can publish and consume open data. The real prize of open data is the innovation that results.

Commissioned by Omidyar Network, this report provides an economic estimate of the value of open data. Whilst exact estimates are bound to remain difficult, the breadth and depth of value highlighted in this report is dramatic. It should provide a compelling incentive to politicians and policy-makers in a time of fiscal austerity. It is a road map for extending open data and reaping the rewards on offer.



SIR NIGEL SHADBOLT  
CO-FOUNDER OF THE OPEN DATA INSTITUTE

## Overview

### AIMS AND SCOPE

This study quantifies and illustrates the potential of open data to the G20 and Australian economies estimating what proportion of the G20's 2% growth target can be delivered by a vigorous open data policy agenda.

Open data embraces a number of data sources and types:

- *Government data*, or public sector information (PSI) is central. It includes all data collected or funded by government agencies at all levels (e.g. national statistics, operational data from service delivery).
- *Research or science data*, especially from publicly funded research. A high-profile example of open science data is the Human Genome Project, but much day-to-day research data remains hidden from view.
- *Private sector data* (e.g. seismic data from mineral exploration) that, with appropriate incentives and privacy protections, may be released.

The estimates in the report presume strong and effective privacy safeguards. Lateral Economics takes the view that any endeavour to seize the opportunities of open data to improve lives will founder if it is not built on the imperative of governments retaining citizens' trust that their privacy will not be compromised.

While individual countries have strong interests in pursuing their own open data agendas, these individual efforts can be greatly strengthened through international cooperation through institutions such as the G20. Since data takes its meaning within standards which permit comparisons, whole new areas of international open data opportunity can be unleashed through standardisation and harmonisation of policies, practices and standards across jurisdictions. (e.g. international trade data, see Annex 1).

### THE OPEN DATA POLICY CONTEXT AND OPPORTUNITIES

The most direct policy action involves releasing, and making more accessible, existing government and research data.

This is consistent with agreed international guidelines<sup>1</sup> and initiatives<sup>2</sup> emphasising that government and research data be:

- Freely available at no cost, or at the marginal cost of distribution;
- Easily discoverable and readily accessible; and
- Open to unrestricted use and re-use on conditions that are readily discoverable and understandable.<sup>3</sup>

Nevertheless, high-level international statements of good practice or intent and even statements of high government policy must often run the gauntlet of the numerous details and complications that must be worked through by officials. As this report shows, much remains to be done.

### **But there are many further opportunities for government action.**

With the increasing ease with which data can be stored, used and distributed, the marginal benefit of all such activities increases. Governments can enrich current data collections, collect additional data for release, and modify regulations pertaining to information (e.g. standards) so as to facilitate greater use of open data and review regulation in specific industries.

Governments may also create value by modifying the environment under which open data are being used. For example, governments may review the regulatory environment if it obstructs open data, and facilitate the emergence of information standards, which enhance the value of private data – for instance, by facilitating comparisons. This will be more effective when harmonised across jurisdictions, giving international forums such as the G20 particular relevance in fully realising the open data opportunity.

---

1 The OECD Recommendation on Public Sector Information, adopted in April 2008.

2 In January 2004, the then 30 OECD countries, together with China, Israel, Russia and South Africa adopted a Declaration on Access to Research Data. See <http://www.oecd.org/dataoecd/41/52/44384673.pdf>.

3 For a full statement of principles, see OAIC (2011) *Principles on open public sector information*, OAIC, Canberra. Available [http://www.oaic.gov.au/infopolicy-portal/reports\\_infopolicy.html](http://www.oaic.gov.au/infopolicy-portal/reports_infopolicy.html).

## The value of open data

### **OPEN DATA CREATES VALUE BY:**

- Reducing the cost of existing government and private services (enhancing efficiency, doing more with less);
- Enabling new services and improved quality of services (enhancing innovation);
- Improving transparency and accountability (enhancing consumer empowerment and governance); and these effects contribute to
- Engendering greater trust in government, which engenders further benefits, such as greater participation.

Estimating the economic value of all this is challenging. The benefits can include direct values (e.g. investment value and market value), as well as indirect values (e.g. wider social and economic impacts from data use). We examine some of the existing research and evidence, and provide new estimates of the value of open data for Australia and the G20, as well as some case-study-based estimates for major G20 themes.

### **Box 1 International studies of the value of government data**

A handful of widely cited quantitative studies demonstrate the difficulties in estimating the value of open data, with very different estimates arising from different approaches.

Top-down macro-economic approaches risk overestimating the benefits of open data, not least because they typically ignore substitution possibilities for other data to be used instead of open data. But bottom-up approaches built from micro-economic analysis risk underestimation, not least because they tend to miss a wide variety of positive impacts, many of which are serendipitous.

Hence, estimates vary widely. Moreover, economic structures vary from country to country, making direct conversions difficult. Nevertheless, based on scaling from GDP shares, it is possible to get a sense of the value of government data in Australia – with aggregate direct and indirect values of perhaps AUD 25 billion per annum. Similarly, based on GDP shares across the G20 countries, this would suggest aggregate direct and indirect values of perhaps USD 700 billion to USD 950 billion per annum.

We explore two top-down approaches to estimating the value of open data, and a series of case-study examples, in an effort to 'tri-angulate' towards a robust estimate.

### **OPEN GOVERNMENT AND RESEARCH DATA**

This report estimates returns to investment in the creation/collection of data and the impacts of increased data accessibility and (conservatively) focusing on government- and publicly-funded research data only, we provide a likely lower bound estimate of the potential value of open data.



Exploring the additional returns that might arise from doubling data accessibility and use, we find that returns to investment in Australian government and research data might rise by around a present value of AUD 240 billion over 20 years, the mid-point of a range from an increase of AUD 120 billion to AUD 360 billion.<sup>4</sup>

These estimates are no more than indicative. For example, we assume a doubling of accessibility and use, but they could increase by a factor of three, five or even more. With the G20 seeking to increase GDP growth by 2% in the next five years,<sup>5</sup> under our conservative modelling, doubling accessibility and use would add 0.27% of Australia's cumulative GDP over the next five years thus contributing around one-seventh of the G20's target for Australia.

### **BROADENING THE NET TO INCLUDE ALL FORMS OF OPEN DATA**

However the open data agenda goes beyond the domain of more open government and research data. A recent McKinsey Global Institute study suggests that the wider open data agenda – which applies to business data as well as government data – has the potential to unlock approximately USD 3.2 trillion in additional value annually across seven global economic domains.

We began this study thinking that the McKinsey number was probably over-optimistic and should be taken as providing an upper bound estimate. However, for at least two reasons, we came to think of the estimate as possibly conservative.

- Firstly, traditional methodologies often underestimate the impact of general-purpose innovations which typically revolutionise life in ways scarcely contemplated at their birth. Further, more formally theoretical considerations suggest non-linear benefits from opening data resources for use, again suggesting possibly large surprises on the upside. Thus, for instance, we provide an example of new digital approaches improving the productivity of some important tasks by 10,000%. (See §3.1.2).
- Secondly, our own case study based estimates, on which we have sought to err on the side of conservatism suggest similar magnitudes to the McKinsey study. And yet they are only a snapshot of the G20 themes, which themselves focus on a sub-set of G20 economic activity.

#### *Australian impacts*

Although McKinsey did not estimate Australia's share of the open data dividend, simple apportionment of the McKinsey figure to the relative size of the Australian economy puts the total potential value of open data to Australia at around AUD 64 billion per annum. (See Figure 1).<sup>7</sup>

---

4 The return from a single year's data collection spending would be around AUD 17 billion over 20 years (NPV) – effectively a return 1.5 times greater than the investment.

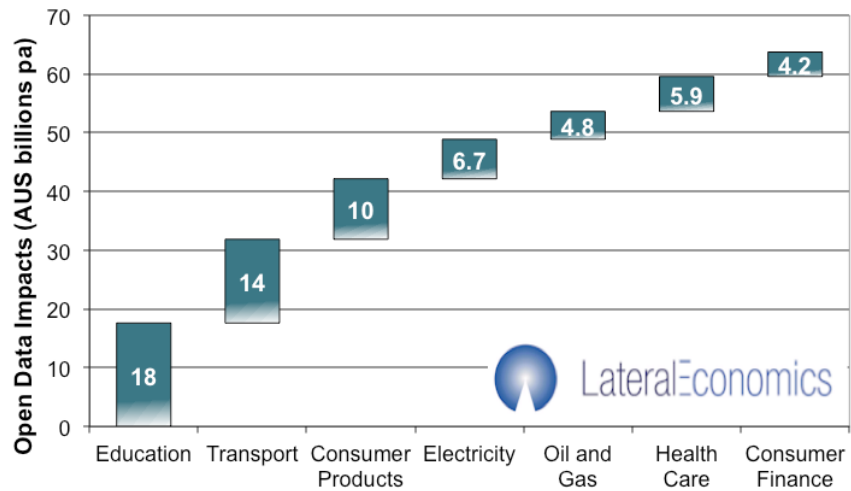
A 20-year period is used to capture the value of data that has a multi-year 'shelf-life', with returns to investment accruing over a number of years. The range of rates of return used is derived from the literature investigating the social rate of return to R&D.

5 Communiqué, Meeting of Finance Ministers and Central Bank Governors, Sydney, 22-23 February 2014.

6 McKinsey Global Institute (2013) Open data: Unlocking innovation and performance with liquid information. New York.

7 Note: Extrapolation of McKinsey (2013) USD estimates to Australia and G20 based on World Bank 2012 USD GDP in current prices: <http://databank.worldbank.org/data/views/reports/tableview.aspx?isshared=true> (Extracted April 10th, 2014). Conversion to AUD at parity.

**Figure 1 Value of open data for Australia (AUD billions per annum)**



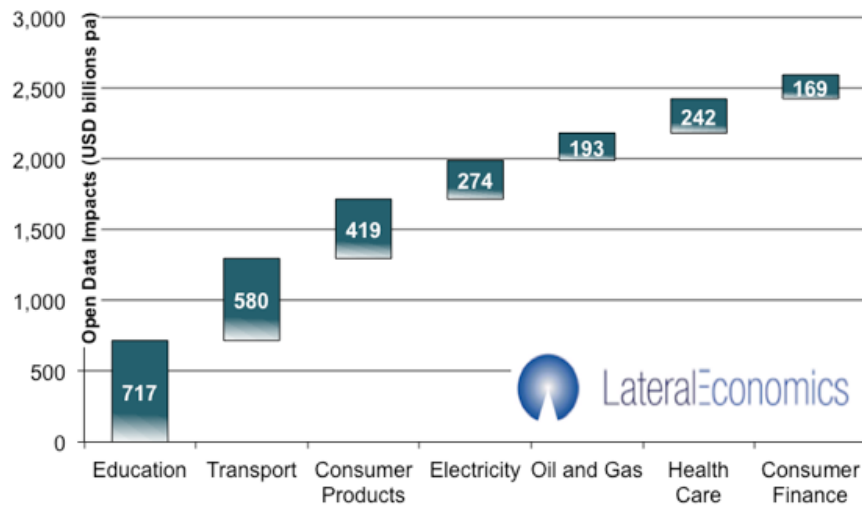
Source: Based on McKinsey Global Institute (2013) *Open data: Unlocking innovation and performance with liquid information*. New York.

Given that the impact identified is output rather than value added and that McKinsey does not identify the share attributable to open data policies, we scale these impacts to estimate their contribution to GDP. We conservatively assume that two-thirds of the impact is value-added, half of this is attributable to open data policies, and three-quarters remains to be realised. This produces a contribution of reinvigorated open data policies to Australia's cumulative GDP growth of AUD 16 billion per annum or around 1% of GDP over the next five years or half the G20 2% growth target.

**IMPACTS AND THE G20 GROWTH AGENDA**

Based on the G20 countries' aggregate share of global GDP, the potential value of open data to the G20 would be around USD 2.6 trillion per annum. Figure 2 provides an approximate breakdown.

**Figure 2 Value of open data for G20 countries (USD billions per annum)**



Source: Based on McKinsey Global Institute estimates, pro rata on GDP shares. Authors' analysis.

Using the same approach to illustrate the effects of open data, we find that G20 output could have increased by around USD 13 trillion cumulatively over the next five years. Again, assuming that two-thirds of the impact is value added, one-half is attributable to open data policies, but that *four-fifths* remains to be realised,<sup>8</sup> open data would make a contribution to aggregate G20 countries' cumulative GDP of around 1.1% over the next five years or 55% of the G20's 2% additional growth target.

## REALISING THE OPEN DATA OPPORTUNITY

Governments should intensify the release of existing public sector data – both government and publicly funded research data.

However, much more can be done to promote open data than more vigorous release of government and research data.

In appropriate circumstances, governments can mandate public disclosure of private information (e.g. in corporate reporting on financial and other matters). Likewise, governments can influence information architectures inducing greater data release. Thus, in consultation with industry, mining exploration/development rights might be issued on condition of with expectations greater public disclosure of the resulting seismic data with large productivity gains. The terms of this could be negotiated with the mining sector generally because it has a collective interest in sharing data, even if no individual miners have such an interest in sharing their data unilaterally without others following suit. Government has similar leverage in areas in which it delivers or funds service delivery (e.g. in health, education and elsewhere).

Government leadership can also facilitate the emergence of conditions that stimulate the private release of information. It is easy to see why firms might not want to publish data on their own performance when that performance is bad, but why don't more firms publish data when their performance is good? Often the reason is that if they do, it has little market impact because it is impossible to use it to compare performance among firms or with industry averages. The obstacle to greater data openness is not necessarily a reticence to publish so much as the lack of a well-understood public standard against which to report. As this report outlines, even without compulsion, governments can play an important role in facilitating the emergence of such standards with leadership and suasion.

By way of illustration of several of these themes, Annex 5 shows how governments could facilitate the surfacing data on employee job satisfaction. Releasing such data on government employees would generate benefits on its own – helping people identify the best-matched government workplaces to their skills and aspirations. This would improve job matching, job satisfaction and productivity. However, such data release might also be done as part of a larger initiative in which leading employers were invited to collaborate with government on the development of a standard against which they could voluntarily report and so benefit from the demonstration of their superiority as employers. This could generate large, economy-wide productivity gains.

---

8 Given the proportion of the G20's economies that are not very advanced on the open data agenda, we judge the potential gains in the G20 to be somewhat higher than they are in Australia though this is balanced by the fact that two large G20 economies, the UK and the US are more advanced than Australia in their implementation of open data (See §2.2.3).

## Box 2 The G20 and the emergence of information standards

Governments can also 'seed' open information architectures with strategic donations to the commons, which help facilitate the emergence of standards. Because information travels in standards, the emergence of a standard may be crucial to the value of information giving it *comparative* significance, which is often crucial.

Secondly, once a standard has emerged, its facilitation of comparisons unleashes *incentives for good performers to report against it*, strengthening the use of the standard and drawing data previously privately held, into the public domain. (See Box 9).

Numerous opportunities also exist for public private partnerships in data (including 'big data') as recognised by OECD and Australian Government policy.<sup>9</sup> Such partnerships range from relatively prosaic ones to deep integration between public and private sectors.

Thus the personal genomic services (as, for instance, provided by 23andMe) could be hugely amplified for the good of the whole health system in a public private partnership. Where 23andMe has acquired around 600,000 customers from a target market of more than one billion, Gruen (2013)<sup>10</sup> proposes that the health system 'bulk-bill' or fund the service and 'nudge' people to opt in, and that, as a condition of public partnership, the genotypic data be available (though with strict safeguards of privacy) as a research asset. This public-private partnership structure could dramatically increase consumer uptake and thus the power and productivity of the resulting research database. Further work may be necessary to address concerns around de-anonymisation of personal data and/or the misuse of any personal data.

Policy can also seek to encourage the emergence of an information architecture that attracts donations to the data commons by assembling data on a platform on which data resources can accumulate to be mashed up against each other and have value added to them by special purpose applications.

It is on this basis that the Sense-t project is funded by the Federal and Tasmanian Governments. It integrates government and private sensor and other environmental data around Tasmania into an open platform for the benefit of all users and on which special purpose apps can be built. Similar platforms could be built around data as a public good envelope around other domains.

9 See Robb, A and Turnbull, M, 2013. *The Coalition's Policy for E-Government and the Digital Economy*, September, at p. 21. <http://paweb-static.s3.amazonaws.com/Coalition%27s%20Policy%20for%20E-Government%20and%20the%20Digital%20Economy.pdf>. See also AGIMO's recent blog post on the same subject at <http://www.finance.gov.au/blog/2014/13/03/2014-seeking-proposals-for-joint-private-public-projects-using-big-data/>

10 Gruen, 2013 "Place for public-private partnerships in cyberspace" *Sydney Morning Herald*, March 29 available at <http://www.smh.com.au/business/place-for-public-private-partnerships-in-cyberspace-20130328-2gxdp.html#ixzz31OOxoBAG>

In many areas progress will be stymied until data architectures are built around users' control of their data.<sup>11</sup> Built on citizens' legislated right to access data corporations' hold on them, the UK government's "Midata" strategy involves a government partnership with corporations in the release of private data to citizens who can then release it on terms they choose, to enable value addition to the data around citizens' needs and interests.<sup>12</sup> This seems a very promising approach, although events in the UK demonstrate the need for governments to retain citizen trust by better understanding and communicating the risks around sharing personal data and ensuring that citizens feel properly informed and confident that their privacy will continue to be protected.

### THE OPEN DATA OPPORTUNITY AND THE G20 AGENDA

The open data opportunity includes an important role for international forums, such as the G20. Our case studies show both the potential impacts of open data, and reveal much about the issues, barriers and policy-related opportunities that arise both at the national and international levels. Figure 3 illustrates their alignment against the themes of the G20's Agenda for Growth and Resilience.

**Figure 3 A G20 Agenda for Growth and Resilience in 2014**

CASE STUDIES	Strong, Sustainable and Balanced Growth			CASE STUDIES
ANNEX 1 Trade Data International	Trade	Boosting Growth through the Private Sector	Investment and Infrastructure	ANNEX 7 Roads Water
	Development		Fiscal and Monetary Policy	ANNEX 3 Managing Economic Cycles
ANNEX 5 Windows on Workplaces, APS Jobs	Employment	Building Global Economic Resilience	Tax	
ANNEX 4 Anti-Corruption	Anti-Corruption		Reforming Global Institutions	
ANNEX 6 Energy	Energy		Financial Regulations	ANNEX 2 Sirca (Tick Data)

Source: *Lateral Economics Analysis*

11 See eg. Big Data: Seizing Opportunities, Preserving Values, 2014, Executive Office of the President, May 1. [http://www.whitehouse.gov/sites/default/files/docs/big\\_data\\_privacy\\_report\\_may\\_1\\_2014.pdf](http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf)

12 <https://www.gov.uk/government/news/the-midata-vision-of-consumer-empowerment>

### **G20 Agenda Item – Annex 1 – Trade**

Exploring the case of open trade data, an expansion of world merchandise trade by just one-tenth of 1% would be worth USD 40 billion per annum. Yet, obtaining and using trade data from multiple jurisdictions is difficult. Even small access fees, the niceties of specific licences, or non-machine readable formats can all involve large transaction costs. How much easier if all such data were harmonised and compliant with high standards of openness not just nationally, but across the G20. The data that each country generates is a potential global public good, as each party to improved trading patterns gains. Open data policies should be harmonised across the G20 countries and beyond. Given empirical evidence on the costs of information barriers to trade, we assume a plausible potential benefit of open-trade data to be 0.6% of merchandise trade, which for Australia would be equivalent to around AUD 1.6 billion or 0.1% of GDP.

### **G20 Agenda Item – Annex 3 – Fiscal and monetary policy**

Open data could help us better manage economic cycles. We think a reasonable estimate of the potential gains is AUD 3.6 billion per annum to Australia's economy alone or 0.23% of GDP.

- Governments possess rich real time data that, even today, is not open or even routinely accessed by government macro-economic managers.
- They use sophisticated models to forecast and assess alternative policy settings. The US Federal Reserve and Australian Treasury are beginning to release open documentation of these models, which will improve public discussion – and the quality of the models. All G20 countries should follow this lead.
- Further, with progressively more software services being delivered from 'the cloud', service providers can aggregate the data into real-time snapshots of the macro-economy. Governments could seek out arrangements with them to supply anonymised data to 'take the pulse' of the economy in real time.

Because the damage done by economic downturns rises approximately as the square of their severity, even very small improvements generate larger relative benefits to the economy.

### **G20 Agenda Item – Annex 4 – Anti-Corruption**

Corruption is a global issue, affecting developed and developing countries. The European Commission estimates the cost of EU corruption to be around 1% of GDP, equivalent to around EUR 120 billion per annum, though it is very difficult to measure the indirect costs of corruption.<sup>13</sup> There is now a growing body of evidence that open data plays a role in reducing corruption.<sup>14</sup> Open data can reduce the extent of corruption by both reducing its private returns and making it easier to detect. Based on the evidence, we think it reasonable to suggest that the costs of corruption would be reduced by of the order of 10%. If so this would produce benefits to the Australian economy of around AUD 1.5 billion per annum or 0.1% of GDP.

Accessing these benefits requires action to both release data and make data interoperable; to be combined and compared with other data. To this end, governments should go beyond transparency and not stop at just releasing the data. They should also collaborate on technical standards and ensuring data meets the needs of data users both domestically and internationally.

---

13 The costs of corruption are not easily measured. The cited figure is based on estimates made by other institutions which suggest that costs of corruption amount to 5% of GDP worldwide.

14 For instance, open data from the Canada Revenue Agency helped to expose one of the biggest tax frauds in Canada's history. Eaves, David (2010) Case Study: How open data saved Canada \$3.2 Billion, Available <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>.

### ***G20 Agenda Item – Annex 5 – Employment***

Open data could transform labour markets. Higher quality information on employment conditions would facilitate better matching of employees to organizations, producing greater job-satisfaction and improved productivity. A longer-term benefit would be improved workplace quality as firms sought to improve workplace conditions in response to this increased transparency. This, in turn, should encourage greater work-force participation and, for those countries where such measures were common, improved international mobility of skilled workers. A very partial calculation of these benefits from the literature that exists suggests gains of around AUD 3.4 billion per annum over the first five years or 0.22% of GDP. This counts just the benefits from lower workplace stress and improved job matching of teachers. Because of the partiality of these calculations, the gains are likely to be far greater.

### ***G20 Agenda Item – Annex 6 – Energy***

The energy sector (incorporating the extraction and supply of energy), is a significant sector in all G20 economies and particularly Australia. Annex 6 explores the potential contribution of open data to the energy sector, particularly in the areas of generational and distributional optimisation, infrastructure design and development, and energy use efficiency.

Given the significant size of the energy sector, a small improvement in efficiency can have a large impact. Many opportunities have been identified in using open data to reduce the cost of energy supply and improve energy efficiency. These opportunities will provide improved information for resource exploration and development, investment decisions, efficient operations management and energy efficiency by consumers. The benefits to an open data agenda also include the possibility that greater transparency will lead to a better understanding of regulatory costs and benefits and be a catalyst for regulatory reforms.

Most nations share common energy efficiency issues, and developments by one nation can provide the lead for others. A conservative estimate of a 2% improvement in efficiency in the sector, which generates less than half the McKinsey estimate, would add 0.11% to GDP, or around AUD 1.7 billion to Australia's economy annually.

### ***G20 Agenda Item – Annex 7 – Infrastructure***

Open data can assist with both maximising the value obtained from existing infrastructure, and ensuring the efficient development of new infrastructure (Annex 7). We conservatively estimate these two benefits at AUD 3.6 billion per annum (comprising, after rounding, AUD 2 billion in more efficient use and AUD 1.5 billion in more efficient development).

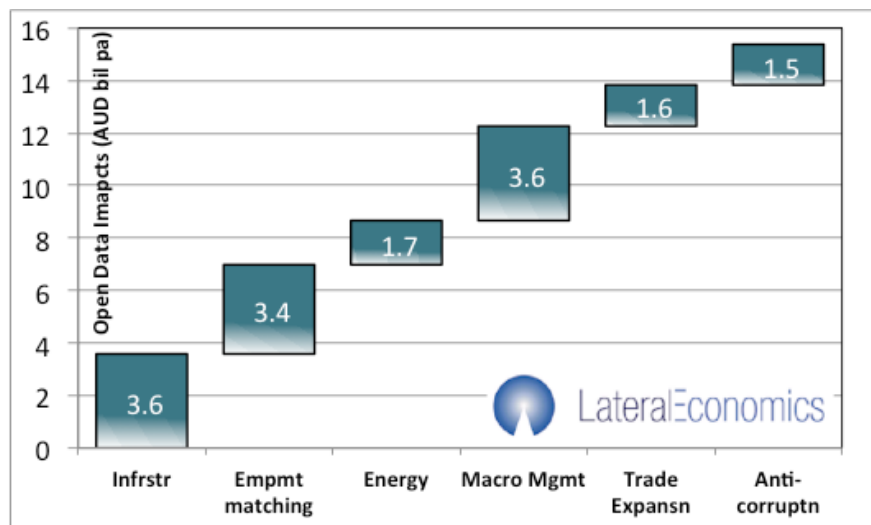
Open data can help improve utilisation in numerous ways; in using real-time information to manage congestion, in providing services associated with the infrastructure and to facilitate regulatory reform (e.g. congestion pricing). It can also improve implementation planning and coordination, and procurement processes as well as better public scrutiny and so governance of investment decisions. Current infrastructure asset information is fragmented and inefficient. Exposing current asset data would be a significant first step in understanding gaps and providing new insights. Doing so in a consistent and harmonised way across G20 countries could provide a wealth of comparative information, enabling greater efficiency in infrastructure development and use.

## **THE OPEN DATA OPPORTUNITY ILLUSTRATED THROUGH LATERAL ECONOMICS' CASE STUDIES**

Figure 4 below illustrates the magnitude of the gains available from a range of case studies. We believe our assumptions in these case studies are conservative, and in some cases, for instance regarding workplace benefits dramatically so. Nevertheless they suggest that the gains available from addressing just a sub-set of the G20 themes – which themselves are just a sub-set of the domains over which policy can accelerate

growth – is over AUD 15 billion or 1.0% of GDP, which is strikingly similar to the numbers produced by our scaling of the McKinsey Global Institute results.

**Figure 4 Value of open data for Australia (AUD billions per annum)**



Source: Lateral Economics Analysis

## CONCLUSION

In addition to its potential to contribute half of the G20's 2% growth target in Australia and slightly more on average within the G20 partners, there is an additional hugely compelling argument for open data. Unusually amongst policies that make a substantial contribution to economic growth it does so with minimal creation of 'losers' from the policy. Virtually any comparable micro-economic reform would involve hard choices and cognate political difficulties for those introducing the change. By contrast, if governments proceed prudently – for instance, building their initiatives on ensuring adequate and well understood privacy protection – the kinds of changes envisaged in this report are all things that would be welcomed by the community because of the obvious good that they would achieve. They provide opportunities for gain without pain.



## 1. Introduction

The open data agenda has an irrefutable commonsense to it. Data, the collection and curation of which is funded and/or conducted by governments is, in the words of Australia's Freedom of Information Act, a "national resource". It follows, from both economic efficiency and civic equity that, as the act specifies, it should be "managed for public purposes" or to be more specific, so as to maximise the public value of the public resource.

Because information is so easily reused and re-purposed, and because it is often difficult for one agency – such as the one funding, collecting and/or curating the data – to even fully understand the best uses for it, there is a strong presumption that such data should be open. That is, it should be open to all comers, in a machine-readable form, to enable all comers to make use of the data and to add value to it however they wish. Whole industries have grown up around the infrastructure or 'platform' that is government funded/ provided data – such as those industries adding value to meteorological and/or geographical data – and the more open the original data is made, the more value can be added.

However, progress on the open government data agenda has been slow. Governments have typically signed on to the 'open data' agenda with various high-level announcements and even detailed policy work. Yet agencies and those within them find ways to delay openness to minimise institutional risks. Those risks range from legitimate ones which should be addressed – for instance, privacy, security and confidentiality – to darker, unacknowledged motives such as avoiding the accountability that open data can bring on. Further, as the private sector moves into the sunlit uplands of 'big data', governments are being left behind and the value their data could have 'in the wild' continues to go untapped.

In important ways, the open data agenda could be taken much further. It could be taken further to seize the opportunities of bringing more research and privately collected data into the open – as so much privately written code has now been brought into the public domain by being incorporated into open source software projects. Further, open data should stand for more than simply surfacing what data is funded/collected by government. It should stand for a deeper commitment to ongoing optimisation of government data collection, to more felicitous partnership with both for profit enterprise and civil society in resourcing and curating data collection and ultimately to an optimal evolution of our economy and society as one rich in knowledge, where decisions are more felicitous because they are better informed. The gains from these broader possibilities are immense.

### Box 3 Public goods of necessity and opportunity and the open data agenda

Since Ostrom and Ostrom (1977), the standard textbook treatment defines such goods as 'non-excludable' and 'non-rival'.<sup>15</sup> Non-excludability creates a free-rider problem, as once provided, the public good's benefits can be shared whether or not users pay its price. Yet non-rivalrousness generates a free-rider opportunity. Once provided, one wants to maximise use of a non-rival good.

Economists have been more anxious about public goods free-rider problems than the opportunities. Thus the Ostroms (1977) teach us, canonically, that public goods' inability to command market prices – "presents serious problems in human organization". Their focus is on all those public goods that cannot be adequately funded privately and which accordingly require government provision, funding or regulation – such things as suburban roads, parks, clean air and defence. Such 'text-book' public goods might be called public goods of necessity.

Yet a growing class of public goods are in fact technically excludable, but entail such free-rider opportunities that they are provided freely. Technically excludable, Wikipedia could operate behind a pay wall, but philanthropic funding underwrites the maximisation of its social value as a free public good.

Thomas Jefferson put the case for ideas as public goods of opportunity more forcefully than any politician has since: "He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me." The open data agenda is similarly focused around the notion that there is a vast trove of value from maximising the extent to which data resources become public goods of opportunity. Where governments already generate such data, then in the absence of privacy or security concerns, the case for making such goods public goods of opportunity is overwhelming.

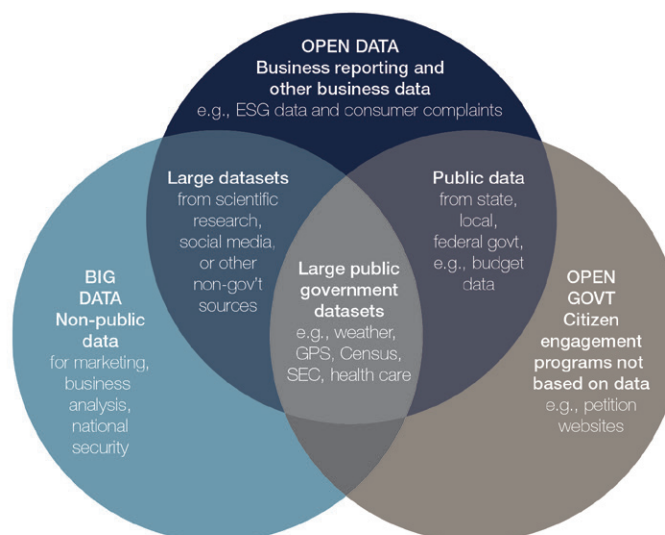
#### 1.1 Objectives and scope

The aim of this study is to quantify and illustrate the potential value of open data to the G20, including the Australian economy. The goal is to provide an evidence base for the Australian Government and other G20 governments, demonstrating the importance of open data as a policy focus that can contribute to the G20 agenda, most particularly its growth agenda to increase G20 growth in the next five years by 2%.<sup>16</sup> For this reason, the study seeks to estimate what proportion of that growth agenda might be delivered by the implementation of a vigorous open data agenda.

15 Ostrom, V. & Ostrom, E. (1977) 'Public goods and public choices,' In E.S. Savas (ed.) *Alternatives for delivering public services: towards improved performance*, Westview Press, Boulder, pp.7-49.

16 [https://www.g20.org/g20\\_priorities/g20\\_2014\\_agenda](https://www.g20.org/g20_priorities/g20_2014_agenda).

**Figure 5** Types of open data



Source: From Gurin, J. (2014) *Open Data Now*, McGraw Hill. Available <http://www.mcgrawhillprofessionalbusinessblog.com/2014/02/18/an-infographic-big-data-is-big-open-data-is-revolutionary/>.

There are a number of data types involved in the open data agenda (Figure 5).

- Primary among them is government data or Public Sector Information (PSI), which includes a wide range of data collected or funded by national, regional and local governments and government agencies that can be purposefully collected (e.g. national statistics, meteorological, mapping and other spatial data) or arise as an integral part of the government function (e.g. business registration, court records).
- Another major area is research or science data, especially that arising from research that is publicly funded.
- There is also an enormous range of private sector data that could be open for both public and private benefit (e.g. vehicle tracking information for traffic management and infrastructure design and development, barcode sales data for economic management, such as estimation of consumer price index).

The estimates in the report presume strong and effective privacy safeguards. Lateral Economics takes the view that any endeavour to seize the opportunities of open data to improve lives will founder if it is not built on the imperative of governments retaining citizens' trust that their privacy will not be compromised.

The open data agenda is broad, with very different levels of policy attention and implementation progress across these data types.

## 1.2 Outline of this report

The next section (Section 2) presents a brief review of the international and local Australian policy context focusing primarily on open government data and open science data.

Section 3 explores the economic value of open data, examining major studies from around the world and providing some indicative estimates of the value of open data for Australia and beyond. Case studies characterising the impacts and value of open data in the G20 Agenda theme area are discussed in summary throughout and a number of more sustained case studies are documented in Annexes 1-7.

Section 4 outlines the way forward from where we are. Taking account of the fact that Australia has been a world leader in open data, the chapter sets out the case for intensifying the campaign to release existing data assets and then sketches out a landscape of more ambitious measures to more fully grasp the economic and social benefits that can be unleashed through open data.

## 2. Policy context

Over the last decade there has been increasing awareness of the benefits of more open access to data, particularly government data or Public Sector Information (PSI) and science (research) data. That awareness is based on economic principles and evidence, and it finds expression in policy at institutional, national and international levels.

### 2.1 Economic principles

Information has public good characteristics. It is 'non-rivalrous' – one person's consumption of it doesn't prevent others consuming it. And it is often non-excludable – even if it is behind a 'paywall' it is difficult to prevent information spreading to others who need not pay. In general, the private sector will tend to under-produce such goods as it is difficult to realise their full value. This justifies public sector supply of information (Nilsen 2007; 2009).<sup>17</sup> Further, Stiglitz et al. (2000) concluded that the theoretical underpinnings of the private versus public trade-off shifts as the economy moves online with a larger public role in the digital economy.<sup>18</sup>

Economists argue that pricing above the marginal cost of dissemination is inefficient because it results in a deadweight loss and eliminates some of the consumer surplus – some people will be prevented from enjoying the benefit of the good even though their consumption of the good would come at little or no marginal cost to the producer. For most government data, marginal cost pricing will be the preferable option (Pollock 2009, p40).<sup>19</sup> The marginal cost of online dissemination is close to zero. Cost recovery through such mechanisms as user fees is never welfare enhancing (Nilsen 2007; 2009).<sup>20</sup>

Some public sector agencies retain copyright on the information they produce to enable them to control access and the conditions of use, so that they can impose a toll on that access and use and prevent uses that would result in the loss of control. However, analysts have concluded that government or "Crown" copyright has social costs and a negative economic impact (Nilsen 2009).<sup>21</sup> Moreover, if the price is set to the marginal cost of online dissemination (zero), there is no purpose to be served by retaining copyright. Indeed, as the costs of disseminating and accessing information have declined, the transaction costs associated with charging for access to information have come to constitute a major barrier to access in themselves (Quiggin 2010).<sup>22</sup> Studies of transaction costs relating to information suggest that they are significant (Poplin [nee Krek]

17 Nilsen, K. (2007) *Economic Theory as it Applies to Statistics Canada: A Review of the Literature*, Statistics Canada, Ottawa. Available [http://www.epsiplatform.eu/sites/default/files/ezpublish\\_media/Nilsen%20Economics%20Paper%202007%20final%20version.pdf](http://www.epsiplatform.eu/sites/default/files/ezpublish_media/Nilsen%20Economics%20Paper%202007%20final%20version.pdf).

18 Stiglitz, J.E., Orszag, P.R. and Orszag, J.M. (2000) *The Role of Government in a Digital Age*, Computer and Communications Industry Association, Washington, D.C. Available <http://www.dol.gov/ebsa/pdf/ccia.pdf>.

19 Pollock, R. (2009) *The Economics of Public Sector Information*, CWPE 0920, University of Cambridge, Cambridge. Available <http://www.econ.cam.ac.uk/dae/repec/cam/pdf/cwpe0920.pdf>.

20 Nilsen, K. (2007) *Economic Theory as it Applies to Statistics Canada: A Review of the Literature*, Statistics Canada, Ottawa. Available [http://www.epsiplatform.eu/sites/default/files/ezpublish\\_media/Nilsen%20Economics%20Paper%202007%20final%20version.pdf](http://www.epsiplatform.eu/sites/default/files/ezpublish_media/Nilsen%20Economics%20Paper%202007%20final%20version.pdf).

21 Nilsen, K. (2009) 'Enhancing Access to Government Information: Economic Theory as it Applies to Statistics Canada,' in Paul Uhler (ed.) *The socioeconomic effects of public sector information on digital networks: Towards a better understanding of different access and reuse policies*, National Academies Press, Washington DC. Available [http://www.nap.edu/openbook.php?record\\_id=12687&page=40](http://www.nap.edu/openbook.php?record_id=12687&page=40).

22 Quiggin, J. (2010) *The value of Public Sector Information for Cultural Institutions*, Report to the Gov 2.0 Taskforce. Available <http://gov2.net.au/projects/project-6/>

2010).<sup>23</sup> Hence the efficient economic solution for the dissemination of government and publicly funded research data is likely to be free *libre* and free *gratis* (i.e. making it freely available online *and* using unrestricted licensing, such as Creative Commons).

Nevertheless, it is important to be mindful of the enormous diversity of government data producers, data users, and data types, and of the potential dangers of applying a simple one size fits all prescription.<sup>24</sup> Free may not be the only price. For example, governments may pursue a range of pricing strategies for spatial data, from free online and pricing at the marginal cost of distribution, through to full cost recovery.

Cost recovery has some advantages, such as strong price signals helping to focus data creation/collection on in-demand data. It makes data producers independent of direct government funding, and can help to support data quality and frequency of updating. Conversely, it can lead to high licensing and administration costs for both producers and users, and limit take-up and use because of the price barrier. Free online access and pricing at marginal cost promotes use and the realisation of benefits and reduces licensing and administration costs, but it can leave data producers isolated from user needs and dependent on direct government funding.

In general, the benefits arising from free online data outweigh the costs. Moreover, the problems arising can be addressed in other ways than through pricing, such as fostering close relations between producers and users so that 'market signals' can be delivered directly.

## 2.2 Policy response

This chapter considers how governments have and may act to maximise the open data opportunity. The most direct policy action involves releasing and making more accessible existing government and research data. However, there are other opportunities for government action. For example, governments may collect additional data for release, modify regulation pertaining to information (e.g. privacy and standards) so as to facilitate greater use of open data, and review regulation in specific industries so that the full benefits of open data may be realised.

Moreover, there are major opportunities for cooperative and collective action by governments in such forums as the G20. These include:

- The harmonisation of national policies and regimes so as to foster global markets and developments based on open data;
- The standardisation of requirements and formats; and
- The identification of best practices around the world and leadership in the diffusion of such practices through international forums.

Further possibilities are explored in Section 4.

---

23 Poplin, A. (2010) Methodology for Measuring the Demand Geo-information Transaction Costs: Based on Experiments in Berlin, Vienna and Zurich, *International Journal of Spatial Data Infrastructures Research*, 2010, Vol.5, 168-193. Available <http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/download/169/207>.

24 Deloitte (2013) *Market Assessment of Public Sector Information*, Report to the Department for Business, Innovation and Skills, London. Available [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf)

## 2.2.1 Release and accessibility

Defining government data as “including information products and services, generated, created, collected, processed, preserved, maintained, disseminated, or funded by or for the government or public institution,” OECD (2005) explored the economic issues and provided an overview of developments within OECD countries at that time. They concluded that access, pricing and licensing policies make a substantial difference in the levels of access to, use and re-use of government data.<sup>25</sup>

The *OECD Recommendation on Public Sector Information*, adopted in April 2008, provided policy guidelines designed to improve access and increase use of government data through greater transparency, enhanced competition and more competitive pricing.<sup>26</sup> The OECD recommended:

- Maximising the availability of PSI for use and re-use based upon the presumption of openness as the default rule;
- Strengthening awareness of what PSI is available;
- Ensuring methodical data collection and curation practices to enhance quality and reliability; and
- Maximising the integrity and availability of PSI through the use of best practices in information management.

These principles were derived from existing best practice and became the basis for further developments in PSI access regimes in OECD countries and elsewhere around the world.

Similar principles were reflected in the recommendations of the Australian Government 2.0 Taskforce, which concluded that government should make government data open, accessible and reusable, suggesting that by default government data should be: free, based on open standards, easily discoverable, understandable, machine-readable, and freely reusable and transformable. In turn, these principles were embodied in the Australian Declaration of Open Government and given force in the passage of legislation reforming the Freedom of Information (FOI) Act and establishing the Office of the Australian Information Commissioner, the Government's response to the Government 2.0 Taskforce report and its response to the *Ahead of the Game: Blueprint for the Reform of Australian Government Administration* report.<sup>28</sup>

Internationally, G8 countries adopted the G8 Open Data Charter in June 2013, declaring, inter alia, that:

*We, the G8, agree that open data are an untapped resource with huge potential to encourage the building of stronger, more interconnected societies that better meet the needs of our citizens and allow innovation and prosperity to flourish.*

*We therefore agree to follow a set of principles that will be the foundation for access to, and the release and re-use of, data made available by G8 governments. They are: Open Data by Default, Quality and Quantity, Useable by All, Releasing Data for Improved Governance, and Releasing Data for Innovation.*<sup>29</sup>

25 OECD (2005) *Digital Broadband Content: Public Sector Information and Content*, OECD, Paris. Available <http://www.oecd.org/sti/ieconomy/36481524.pdf>.

26 OECD (2008), *OECD Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, OECD, Paris. Available <http://www.oecd.org/dataoecd/41/52/44384673.pdf>.

27 Government 2.0 Taskforce (2010) *Engage: Getting on with Government 2.0*, Department of Finance and Deregulation, Canberra. Available <http://www.finance.gov.au/publications/gov20taskforcereport/index.html>.

28 Australian Government (2010) *Declaration of Open Government*, 16 July 2010, Department of Finance and Deregulation, Canberra. Available <http://www.finance.gov.au/blog/2010/07/16/declaration-open-government/>

29 G8 Open Data Charter, June 2013. Available <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>

Similar policy initiatives have been adopted for research data, focusing on the data produced by publicly funded research. In January 2004, ministers of science and technology of OECD countries met in Paris, a meeting chaired by Australia, to discuss international guidelines on access to research data. At that meeting, the governments of the [then] 30 OECD countries and of China, Israel, Russia and South Africa adopted a Declaration on Access to Research Data from Public Funding, which sought to maximise the return on public investment in research through open science data.<sup>30</sup>

In all cases, the underlying principles are that government and research data should be:

- Freely available at no cost, or no more than the marginal cost of distribution;
- Easily discoverable and readily accessible; and
- Open to unrestricted use and re-use on conditions that are readily discoverable and understandable.<sup>31</sup>

As Eaves (2009) put it: “if it can't be spidered or indexed, it doesn't exist; if it isn't available in open and machine readable format, it can't engage; and if a legal framework doesn't allow it to be re-purposed, it doesn't empower.”<sup>32</sup>

Nevertheless, while guiding policies have been formulated and announced, much remains to be done to implement these policies and fully realise the benefits of open data.

## 2.2.2 Expansion and regulation: The G20 opportunity

One implication of the increasing ease with which open data can be stored, used and distributed is that the marginal benefit of all such activities increases. An appropriate government policy response may, therefore, be investing in enriching the data currently being captured or investing in capturing other data in entirely new domains.

Following the observation of Stiglitz et al. (2000) that the theoretical underpinnings of the private versus public trade-off shifts as the economy moves online, with a larger public role in the digital economy,<sup>33</sup> government (as opposed to private) investment may be appropriate due to the public good nature of data, and/or the advantages governments have in some domains. By the same token, governments have a collective interest in improving the data ecosystem, because the benefits radiate beyond national borders as demonstrated throughout this report.

Governments may also create value by modifying the environment under which open data are being used. For example, governments may review the regulatory environment that may act as a barrier to open data, and facilitate the development of information standards through international forums, such as the G20.

---

30 OECD (2007) *OECD Principles and Guidelines for Access to Research Data from Public Funding*, Paris. Available <http://www.oecd.org/science/sci-tech/38500813.pdf>. See also <http://www.oecd.org/science/sci-tech/sciencetechnologyandinnovationforthe21stcentury/meetingoftheoecdcommitteeforscientificandtechnologicalpolicyatministeriallevel29-30january2004-finalcommuniqu.htm>.

31 For a full statement of principles, see OAIC (2011) *Principles on open public sector information*, OAIC, Canberra. Available [http://www.oaic.gov.au/infopolicy-portal/reports\\_infopolicy.html](http://www.oaic.gov.au/infopolicy-portal/reports_infopolicy.html).

32 Eaves, David (2009). The Three Laws of Open Government Data, Available <http://eaves.ca/2009/09/30/three-law-of-open-government-data/>.

33 Stiglitz, J.E., Orszag, P.R. and Orszag, J.M. (2000) *The Role of Government in a Digital Age*, Computer and Communications Industry Association, Washington, D.C. Available <http://www.dol.gov/ebsa/pdf/ccia.pdf>.



#### **Box 4 Trade data: the value that can be added by international coordination**

TradeData International shows how open data can improve the efficiency of international trade and facilitate economic development in developed and emerging regions as well as the importance of international coordination in open data through institutions such as the G20. Founded in 1995 TradeData collects and analyses global trade data from monthly updated customs returns.

This allows it to provide commercial intelligence and research to government and industry clients like the State Government of Victoria which commissioned TradeData to analyse the state's industry strengths against detailed analysis of the prices at which Victoria's exports could be landed in various global ports identifying not only with what products, but also where Victoria's producers would be most competitive. This enabled a much more informed base around which to target export promotion and development. TradeData has done similar work for developing countries for instance assisting in identifying opportunities for Pakistan's smallholder farmers.

Over the years, TradeData has encountered various difficulties in providing its services associated with obtaining and using government data from multiple jurisdictions and regions. There can be price barriers, as there is a mix of free and paid data, and the data can require related processing/consulting services from the data providers. TradeData must pay many hundreds of thousands of dollars in data fees not to mention the transaction costs involved in then taking delivery of the data. How much easier if all government data of this kind were freely available?

Varying licensing conditions also present serious problems requiring TradeData to understand and comply with differing licenses from different countries, different legal jurisdictions, in different languages, etc. How much easier if all governments reported to similar standards and used standard and permissive licensing, such as creative commons?

The data that each country generates is a potential global public good as overwhelmingly each party to new trading patterns gains and the data can help the global economy search out its most efficient configurations of production, exchange and consumption. How much easier if open data policies were harmonised across the G20 countries and beyond?

In short, there remain many opportunities where governments can act, unilaterally and collectively to facilitate open data and to encourage an environment that is maximally encouraging of the open data opportunity.

#### **2.2.3 International comparisons of open data policy implementation**

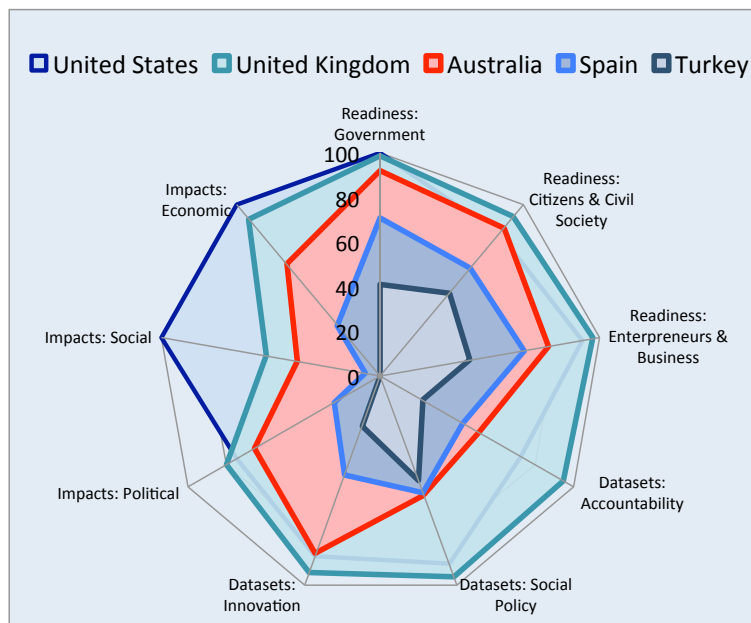
The United Kingdom and the US led the open government/open data movement followed shortly afterwards by Australia. The UK's Power of Information Report in 2007 was a landmark as was the election of Barack Obama in the US who signed the Memorandum on Transparency and Open Government on his first day in office. Australia was a fast follower, and praised at the time for its 2009 Government 2.0 Taskforce. Gartner Managing Vice President for the public sector, Andrea Di Maio, rated the Taskforce the most highly of any specific development in open government in that year.<sup>34</sup>

---

34 [http://blogs.gartner.com/andrea\\_dimai/2009/12/29/a-year-in-review-top-ten-for-government-2-0-in-2009/](http://blogs.gartner.com/andrea_dimai/2009/12/29/a-year-in-review-top-ten-for-government-2-0-in-2009/).

The intervening years have seen progress continue. Though it continues to be ranked as the third in the G20, Australia's progress has probably been somewhat less strong than the UK and US, which have each been vigorous and consistent innovators.

**Figure 6 Relative progress on open government data areas**



Source: <http://theodi.github.io/open-data-barometer-viz/>

**Table 1 G20 Open Data Barometer Rankings, 2013**

Country	Readiness	Implementation	Impact	Overall
United Kingdom	100	100	80	100.00
United States	95	87	100	93.38
Australia	88	65	51	67.68
Canada	79	64	52	65.87
Germany	75	63	54	65.01
France	79	64	39	63.92
Netherlands	86	67	21	63.66
Korea Republic	77	55	25	54.21
Japan	77	47	28	49.17
Spain	67	49	21	48.19
Italy	50	43	46	45.30
Russia Fed.	54	40	49	44.79
Switzerland	65	42	27	43.24
Mexico	49	45	8	40.30
Brazil	66	32	28	36.83
India	57	34	10	33.38
Turkey	42	31	0	27.58
China				N/A*
Indonesia				N/A*
Saudi Arabia				N/A*

\* N/A Not Available.

Source: Open Data Barometer 2013. Available <http://www.opendataresearch.org/content/2013/535/get-data-open-data-barometer-2013>

**Table 2** G20 Open Knowledge Foundation's Open Data Index, 2013

<i>Country</i>	<i>Rank</i>	<i>OD Index Score</i>
United Kingdom	1	940
United States	2	855
Netherlands	3	740
Australia	4	660
Canada	5	590
Italy	6	515
France	7	510
Switzerland	8	480
Spain	9	460
Japan	10	440
Korea Republic	11	425
Russian Federation	11	425
China	12	415
Indonesia	12	415
Germany	13	410
Brazil	13	410
Mexico	14	405
Saudi Arabia	15	280
India	16	215
Turkey		N/A*

\* N/A Not Available.

Source: <https://index.okfn.org/country>.

### 3. The economic value of open data

This chapter explores the economic value of open data, examining major studies from around the world before attempting to provide some estimates of the value of open data for Australia and G20 economies. A series of case studies characterising the impacts and value of open data in the G20 Agenda theme areas can be found in the Annex.

There are a number of ways in which open data can create value. Open data:

- Leads to cost reduction in the provision of existing services both by government and private sector (i.e. doing the same for less cost);
- Enables new services and improved quality of services; and
- Indirectly contributes improvements in government services through improved accountability and engendering greater trust in government.

There are various dimensions of value to consider. Deloitte (2013) note that recent studies have examined three levels of economic value generated by public sector information:

- Direct value (i.e. the revenue generated by government from selling access to public sector information);
- Commercial value (i.e. the revenue generated by private companies through the use of public sector information); and
- Downstream value (i.e. the value to users of products and the wider economic, social and environment benefits generated).<sup>35</sup>

These are progressively more difficult to estimate. Measuring cost reduction is relatively straightforward, using existing cost as a baseline (upper bound), but there is often no obvious upper bound for new benefits. It is difficult to attribute impacts to open data as it is just one part of the equation and there may be many factors contributing to the benefits. In Australia, much government data are already openly available, though much might be done to release more and improve that which is available, and it is difficult to separate existing from potential additional benefits yet to be realised. Ultimately, it is impossible to identify all the benefits arising from open data, as many potential applications have not yet been invented.

Nevertheless, the evidence suggests that government data and science data already play an important role in knowledge-based economies and, with many of the benefits still to be realised, more open data could contribute to further economic growth.

---

35 Deloitte (2013) *Market Assessment of Public Sector Information*, Report to the Department for Business, Innovation and Skills, London. Available [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf)

## Box 5      How Big Data Creates Value

We have identified five broadly applicable ways to leverage big data that offer transformational potential to create value.

*Creating transparency:* In the public sector, for example, making relevant data more readily accessible across otherwise separated departments can sharply reduce search and processing time, increase serendipitous discovery and data 're-purposing' and improve accountability.

*Enabling experimentation to discover needs and improve performance:* IT enables organizations to instrument processes and set up controlled experiments. Using data to analyse variability arising naturally or generated by controlled experiments – and to understand its root causes enabling leaders to manage performance to higher levels.

*Segmenting populations to customise actions:* Big data allows organizations to tailor products and services precisely to market segments' needs. Familiar in marketing and risk management, this approach can be revolutionary elsewhere – for example, in the public sector. Even consumer goods and service companies that have used segmentation for many years are beginning to deploy real-time micro-segmentation of customers to target promotions.

*Algorithmic decision support:* Sophisticated analytics can substantially improve decision making, minimise risks, and unearth valuable insights that would otherwise remain hidden. Such analytics can be deployed to triage highest regulatory and compliance risks, to fine-tune inventories and pricing in response to real-time in-store and online sales. Decisions can be based on the behaviour of whole populations rather than samples. Decision making may never be the same; some organizations are already making better decisions by analysing entire datasets from customers, employees, or even sensors embedded in products.

*Innovating new business models, products and services:* Big data enables companies to create new products and services, enhance existing ones, and invent entirely new business models. Manufacturers are using data obtained from the use of actual products to improve the development of the next generation of products and to create innovative after-sales service offerings. The emergence of real-time location data has created an entirely new set of location-based services from navigation to pricing property and insurance based on car driving habits.

*Source: McKinsey, abridged and adapted by Lateral Economics.*

### 3.1      Introductory digression on non-linearity

The following section explains why we might expect the benefits of open data to be increasing at a possibly accelerating rate with the openness of data and the power of the systems we have to handle data. We do not rely on this possible phenomenon in our report, but outline the issues here to make the conservatism of our own method clear.

Systems and networks can be thought of as collections of individual items and the linkages between them. The best way to think of individual bits of data is as points in such a system with the linkages being its relationship to other bits of data, either directly or as uncovered by the demands of the users. It is clear that each data point gets its value from its contribution to whatever understanding or insight these connections give us. It is always difficult to value outcomes that are, to a significant extent unknown. In such cases, there can be a tendency to talk in terms of black swans or other metaphors and attribute subjective assessments. In the case of open data, however, there is reason to believe that the potential value may go well beyond the direct

benefits that more and better information provide or what can be ascertained by extrapolating existing usage and returns. It may, in fact, be orders of magnitude greater.

This is because a collection of data may have properties that exist at the level of the entire system rather than at the level of the individual bits. In particular, it may share, with many other phenomena ranging from the web, earthquakes, brain functions, traffic jams and the like, the characteristics of a complex network, or what is known as self organized criticality. What this means is, roughly, that once the collection becomes sufficiently large, the entire network is at a critical point. At such a point, any small change, such as an additional input of data, may cause it to rapidly shift to another state.<sup>36</sup>

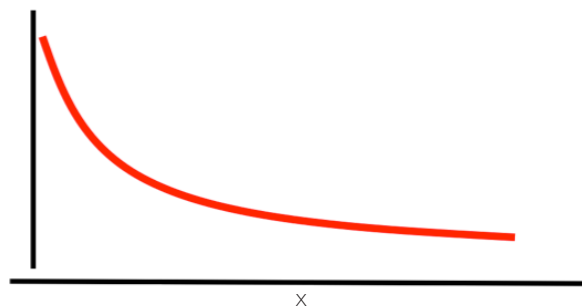
One of the most significant indicators of self-organized criticality in a system is that linkages are scale free and exhibit a power law distribution.<sup>37</sup> A power law distribution means that the fraction of data points with  $x$  connections, or inputs into other collections of data, are expected to follow the function

$$p(x) = bx^{-a}$$

where  $a$  and  $b$  are positive numbers. For example, let  $a = 2$  and  $b = 1$ . Then every bit of data will have one connection, half of the data will have two connections, one hundredth will have ten connections and so on. This function is represented below. The horizontal axis is the number of connections and the vertical is the fraction of the data.

**Figure 7 Power law distribution for data**

$p(x)$



Although a full study is beyond our scope, there are strong reasons for suspecting that open data will exhibit a power law distribution. Three examples are offered:

1. This distribution has been observed across the web, for example, in social media and networking websites, as well as in search engines. Similarly with other data in all their varieties, some bits will rarely be used, say the amount spent on recycling lawn clippings in public parks, whereas other bits of data, such as economic growth rates and GDP, will be more widely used.
2. Most data linkages will follow a process known as preferential attachment. In other words, links between bits of data will not be established at random but will depend on other links. The more useful a bit of data, the more likely it is to be linked. This attachment turns up in wealth distribution,

36 Buchanan, M. (2001). *Ubiquity: Why Catastrophes Happen*, New York, NY: Crown Publishers. And see Wikipedia.

37 Buchanan, *ibid.*

the popularity of web sites, both for visiting and for contributing to, downloads of music and video clips and data searches on the web. For similar reasons, this would be expected in patterns of open data usage. Any distribution with preferential attachment also has the characteristics of a power law distribution.<sup>38</sup>

3. New 'big data' means of tackling knowledge problems are sometimes producing productivity improvements of several orders of magnitude (See the following section).

### 3.1.1 Implications for the value of open data

The fact that open data may exhibit self-organized criticality means that most of the standard valuation methods used in this report may (will probably?) return a figure that is lower than the expected future value, possibly orders of magnitude lower. If open data turns out to be critical, a small increase in the amount of data available should not be treated as simply an increment to the existing network. The important mathematical property of such a system is that any addition always has the potential to cause the entire network to jump to a new state in which the connections and the payoffs change dramatically, perhaps by several orders of magnitude.

Underestimations of this type seem to have been relatively common in relation to information and communication associated activities. Among the examples here are the demand for computers, the effect of mobile phone technology, the time required to sequence DNA and so on. The idea of the runaway process built on power laws is behind the popular idea of 'the singularity'. The idea was put into play by mathematician John von Neumann in 1958, and since paraphrased as the process unleashed at the point that artificial intelligence exceeds human intelligence unleashing a process of "ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue."<sup>39</sup>

An experiment that gives some idea of how criticality works is the button and thread experiment.<sup>40</sup> In the words of Franklin:

*Imagine strewing a multitude of buttons randomly about a bare floor. Now pick two buttons at random and join them by a thread. Put them back. Choose another two, connect and return them. Continue this process, keeping track of the number of buttons in the largest connected cluster, that is, **the largest** group of buttons that could be lifted together by picking up one of its members. Kauffman's computer models of this experiment show that this largest connected cluster grows slowly until the number of threads is a little more than half the number of buttons. Then, suddenly, it grows large very quickly. In models containing 400 buttons, this maximal cluster size goes from under 50 to over 350 as the number of threads rises from just below 200 to just above 250. Plotting a graph of maximal cluster size against number of threads yields a steep S-curve.*<sup>41</sup>

In this experiment, the additional connection has a value, in terms of the number of buttons lifted, well in excess of what we might estimate using more standard extrapolation techniques.

In a similar way, open data always carries with it the probability of such a jump. Moreover, in the case of open

38 Barabási, A.-L.; R. Albert 1999. "Emergence of scaling in random networks". *Science*, 286 (5439): 509–512. arXiv:cond-mat/9910332.

39 "The Singularity," Wikipedia, [http://en.wikipedia.org/wiki/The\\_Singularity](http://en.wikipedia.org/wiki/The_Singularity).

40 Kaufmann, S. 2000. *Investigations*. Oxford.

41 Franklin, Stan, 2000. Review of Stuart Kauffman's, "At Home in the Universe, The search for laws of self-organization and complexity", *Times Literary Supplement*, <http://www.msci.memphis.edu/~franklin/kauffman.html>

data, the jump must always be to a more valued from a less valued state. As with the buttons, more threads can only cause the number lifted to increase. All this is speculative, but it is backgrounded in a well-established body of empirical and mathematical research, which allows estimation of probable returns. In trying to ascertain the value of open data it is always these returns that matter and the probability of large gains from criticality must be properly considered. To err on the side of conservatism, however, they have been omitted from this study so that it is clear that we are not making over-optimistic claims for the benefit of open data.

### 3.1.2 A worked example: Productivity disruption in genomic research

Using genetic information to identify potential health conditions (e.g. disease risks, drug resistance or susceptibility to drug side effects) involves two expenses. The most familiar is the cost tied to sequencing an individual's genetic code. In recent years, this has radically decreased in price from about USD 100 million in 2001 to about USD 5,000 in 2014 for full genome sequencing.<sup>42</sup> Using a genotyping approach, where single spot-differences in DNA are used to build a snapshot of the genome instead of sequencing every nucleotide, reduces the price again, down to below USD 100 per individual.

However, the second major cost has not reduced in line with declining individual sequencing costs; this is the expense associated with gathering the phenotypes (the physical traits, such as diseases), which are identified and linked to specific genetic variations.<sup>43</sup> The traditional approach would be to identify a cohort of people with a condition of interest, usually via medical records. A cohort of about 50 individuals is a minimum starting point, and even this can present difficult and expensive problems for researchers to obtain medical records and contact individuals and meet the inevitable red tape involved in consents and ethics approvals. The cohort is then compared to a control group without the condition, also requiring 50 participants at a minimum. The whole genome of each individual would be sequenced, so that sequencing costs alone would begin at about USD 500,000 without taking into account the work needed to locate potential participants or the subsequent analysis costs.

23andMe have recently demonstrated a different approach where self-reported medical information provided by individuals was used to screen for conditions and build genotype-phenotype maps.<sup>44</sup> These maps were then used to test whether the researchers could replicate the results of other studies that had used more traditional genome-wide phenotyping approaches. The researchers successfully replicated about 70% of the study results they tested. More work is clearly needed, but this is an impressive feat when considering the cost and efficiency savings. The approach adopted by 23andMe used about 20,000 individuals. They were then able to use the self-reported medical data to phenotype 50 conditions at once instead of just one condition.

Using a traditional approach, it would take a small team about twelve months to contact participants, sequence genomes and analyse results. The 23andMe team also took about twelve months, but they were able to examine a much larger set of phenotypes. Put simply, in terms of the conditions that can be phenotyped within one year, the new methodology represents a productivity increase of 5,000%. However, this compares a technology at start-up with all its fixed and learning costs with a mature technology – the incumbent, traditional one. It would be possible to set up the 23andMe database to run such scans on an ongoing basis, and effectively identify patterns as they appear in the data. This takes us somewhere near an infinite increase in productivity!

---

42 National Human Genome Research Institute <https://www.genome.gov/sequencingcosts/>

43 Nature Reviews: Phenomics: The Next Challenge <http://www.nature.com/nrg/journal/v11/n12/abs/nrg2897.html>

44 Tung, Joyce Y., Do, Chuong B., Hinds, David A., Kiefer, Amy K., Macpherson, J. Michael, Chowdry, Arnab B., Francke, Uta, Naughton, Brian T., Mountain, Joanna L. Wojcicki, Anne, Eriksson, Nicholas "PLoS One: Efficient Replication of over 180 Genetic Associations with Self-Reported Medical Data", <http://www.plosone.org/article/info%3A-doi%2F10.1371%2Fjournal.pone.0023473#pone-0023473-t001>.



Similarly, the new methodology offers to almost eliminate the costs of new knowledge of this kind. The cohort of 20,000 people cost USD 2 million to genotype, but now 600,000 customers of 23andMe have paid USD 99 to be genotyped and have donated their own phenotype by completing 23andMe surveys for the range of *private benefits* this brings to them, which includes tapping into existing knowledge about genotype-phenotype associations of interest to them. The public good of more scientific knowledge was thus a by-product of a private investment decision. Given this, we have assumed that the eleven authors of the published study of the work each worked part time on the project doing USD 20,000 worth of work on it at a total cost of USD 220,000. From this, each of the 50 genetic associations was generated at a cost of USD 4,000 each or at 0.89% of the cost of the traditional method, a productivity improvement of 10,000%.

Even if one were to assume that the cost of this work also involved the retail cost of USD 99 for performing all the genomic sequencing of the full 20,000 people involved in the study, the productivity improvement is still over 1,000%.

## 3.2 The evidence to date

Previous studies have focused on different dimensions of value and adopted different approaches and economic methods. Most focus on public sector information (PSI), though some explore research data and private data. This section presents a brief review and analysis of the major studies exploring the economic value of open data, highlighting differences in approaches and the estimates of value.

### 3.2.1 International studies of open government data

There have been a handful of widely cited quantitative studies exploring the value of government data (PSI). These studies demonstrate the difficulties in estimating the value of open government and science data, and that quite different estimates can arise from different approaches. There are also differences between these studies in their scope and foci.

Top-down approaches risk overestimation, not least because they take no account of the possible use of alternative data. Conversely, bottom-up approaches risk underestimation, not least because of difficulties in accounting for wider impacts. Hence, the estimates arising from these approaches vary widely.

#### PIRA

Adopting a top-down approach, PIRA (2000), combined measures of the investment cost (i.e. the amount spent on the collection/generation of the PSI) and expenditure on PSI by users and re-users. Then, for final users, estimated value as expenditure on PSI or, where the PSI was freely available, as the investment cost of its collection/generation; and for intermediaries who add value to the data and re-sell, ascribed a proportion of their value added to the PSI used.<sup>45</sup> A simplifying assumption was that the value of the information accrued in the year that it was collected/generated. This is reasonable for some PSI, such as weather forecasts, but less so for other forms of data, such as mapping. A potential problem with the PIRA approach is that it may overestimate the value of PSI because it does not account for the possible use of alternatives.

By generalising from case studies and scaling up, PIRA (2000), estimated the investment and economic value of PSI in the European Union, putting the former at around EUR 9.5 billion per annum in 1999 and the latter at around EUR 68 billion – equivalent to approximately 1.4% of EU GDP and representing a seven-fold return on investment.

By comparison, PIRA put PSI investment in the United States at EUR 19 billion per annum and economic value at EUR 750 billion – a 39-fold return on investment. They attributed the difference to the US having a

---

45 PIRA (2000) *Commercial exploitation of Europe's public sector information*, European Commission, Brussels. Available [ftp://ftp.cordis.europa.eu/pub/econtent/docs/2000\\_1558\\_en.pdf](ftp://ftp.cordis.europa.eu/pub/econtent/docs/2000_1558_en.pdf)

more open access regime than EU countries, and suggested that the EU could reach US levels if they adopted open data regimes. Importantly, they estimated that EU countries would only need to double the value of PSI for governments to recoup the lost revenues from PSI sales in increased tax receipts.<sup>46</sup>

### MEPSIR

Employing a large survey of PSI producers and users, Dekkers et al. (2006), sought to estimate the size of the PSI market in Europe (the MEPSIR study).<sup>47</sup>

In the MEPSIR study, demand and economic performance were measured in an extensive survey by directly asking both public content holders and re-users for key economic data, such as total turnover against turnover related to PSI, total number of staff against the number of staff dedicated to handling PSI, and estimates of domestic market for a particular type of PSI. The European PSI market value was then estimated from the average revenues multiplied by the average number of re-users per PSI domain, minus the cost of PSI collection/generation. For country estimates, this was distributed according to GDP shares.

Based on the estimates of re-users, they put the overall market for PSI in the EU plus Norway at around EUR 27 billion (approximately 0.25% of aggregated GDP). This is a much lower number than suggested by the PIRA study, despite it being market size rather than value added and coming five years later.

Making some adjustments with the benefit of hindsight, te Velde (2009) suggested that the original estimates were high, perhaps five to nine times too high, and that the value of the PSI market might drop further from EUR 27 to EUR 5 billion or even EUR 3 billion<sup>48</sup> – only around 5% of the PIRA study estimate, and highlighting the enormous range of findings and measurement uncertainties. Our own results suggest something closer to the higher figure.

### DotEcon

DotEcon (2006) produced a report for the UK Office of Fair Trading (OFT) in which they developed a bottom-up approach to estimating the economic value of PSI products and services in the UK, seeing the net economic value of PSI as the willingness to pay for PSI minus the cost of supplying it (essentially, the net consumer surplus). They also looked at the costs (detriment) of barriers to use, including: unduly high prices, distortion of downstream competition, and failure to exploit PSI.<sup>49</sup> While much less subject to over-estimating the value of PSI, a potential weakness of this approach lay in estimating price elasticities of demand, especially where the PSI was supplied free of charge.

While looking at a more limited range of data types, the results indicated that the net value of PSI in the UK was around GBP 590 million per annum in 2005. The costs of the three types of detriment were estimated to be GBP 20 million from high pricing, GBP 140 million from restriction of downstream competition, and GBP

46 The PIRA report was popularised in the widely cited "Borders in Cyber-Space" report: Weiss, P. (2002) *Borders in Cyber-space: Conflicting Public Sector Information Policies and their Economic Impacts*, US Department of Commerce, Washington DC. Available [http://www.nws.noaa.gov/sp/Borders\\_report.pdf](http://www.nws.noaa.gov/sp/Borders_report.pdf).

47 Dekkers, M., Polman, F., te Velde, R. and de Vries, M. (2006) *MEPSIR: Measuring European Public Sector Information Resources*, European Commission, Brussels. Available [http://ec.europa.eu/information\\_society/newsroom/cf/document.cfm?doc\\_id=1197](http://ec.europa.eu/information_society/newsroom/cf/document.cfm?doc_id=1197).

48 Te Velde, R. (2009) 'Public Sector Information: Why Bother?', in Paul Uhler (ed.) *The socioeconomic effects of public sector information on digital networks: Towards a better understanding of different access and reuse policies*, National Academies Press, Washington DC. Available [http://www.nap.edu/openbook.php?record\\_id=12687&page=25](http://www.nap.edu/openbook.php?record_id=12687&page=25).

49 DotEcon (2006) *The commercial use of public information (CUPI)*, Report of861, Office of Fair Trading, London. Available <http://www.opsi.gov.uk/advice/poi/oft-cupi.pdf>.

360 million from failure to exploit PSI – suggesting that the value of PSI could be doubled by resolving the problems identified.<sup>50</sup>

### European Commission

Reviewing the literature on the size and growth of the European market for PSI for the EC, Vickery (2010) noted that the EU27 direct PSI-related market was of the order of EUR 28 billion in 2008.<sup>51</sup> All studies show relatively rapid growth in PSI-related markets and assuming annual growth of 7% the direct PSI-related market would be around EUR 32 billion in 2010.

PSI-related information can be used in a very wide range of direct and indirect applications across the economy and aggregate direct and indirect economic impacts from PSI applications and use across the whole EU27 economy were estimated to be of the order of EUR 140 billion annually, showing clearly that there are large economic benefits from easier access to and greater use of PSI.

These estimates are based on studies of business as usual, but overall economic gains from opening up PSI and providing easy access for free or marginal cost of distribution could be up to EUR 40 billion for the EU27, and aggregate direct and indirect economic benefits could have been of the order of EUR 200 billion in 2008 (43% higher than in the 'business as usual' scenario and equivalent to 1.7% of GDP).

Vickery (2010) went on to note that sectoral estimates of possible gains from the removal of current barriers to access and improving the underlying infrastructure include the geo-spatial sector, where benefits could be increased by some 10% to 40% by improving access, data standards, and building skills and knowledge. Better policies in the area of geospatial applications in local government could help productivity gains almost double over the next five years. Large markets are also estimated in financial, energy and construction sectors.

In terms of efficiency gains in existing operations, improving accessibility of information necessary for obligatory environmental impact assessments could potentially reduce EU27 costs by 20% or around EUR 2 billion per year, open access to R&D results could result in recurring annual gains of around EUR 6 billion per year, and if European citizens each saved as little as two hours per year by more rapid and comprehensive access to public information, this would be worth at least EUR 1.4 billion per year.

### Deloitte/BIS

Building on the bottom-up approach developed by DotEcon, Deloitte (2013) estimated that the value of public sector information to consumers, businesses and the public sector was approximately GBP 1.8 billion in 2011 prices, but suggested that the use and re-use of public sector information has much larger downstream impacts affecting all areas of society beyond the direct customer. On the basis of conservative assumptions, they estimated this could be in excess of GBP 5 billion in 2011 prices. Adding this social value estimate to the calculated value of public sector information to consumers, businesses and the public sector, gave an aggregate estimate of between GBP 6.2 billion and GBP 7.2 billion in 2011 prices.<sup>52</sup>

---

50 Graves, A. (2009) 'The price of everything but the value of nothing,' in Paul Uhrir (ed.) *The socioeconomic effects of public sector information on digital networks: Towards a better understanding of different access and reuse policies*, National Academies Press, Washington DC. Available [http://www.nap.edu/openbook.php?record\\_id=12687&page=40](http://www.nap.edu/openbook.php?record_id=12687&page=40).

51 Vickery, G. (2010) *Review of recent studies on PSI re-use and related market developments*, European Commission, Brussels. Available [http://ec.europa.eu/information\\_society/newsroom/cf/document.cfm?doc\\_id=1093](http://ec.europa.eu/information_society/newsroom/cf/document.cfm?doc_id=1093)

52 Deloitte (2013) *Market Assessment of Public Sector Information*, Report to the Department for Business, Innovation and Skills, London. Available [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf)

### 3.2.2 Australian studies of open government data

Australian studies of PSI markets and the value and impacts of open government data have focused on the major data types, including spatial data, national statistics, and hydrological data. Again, what is being measured, methods adopted, and findings vary.

#### Spatial data

Looking at spatial information in Australia, ACIL Tasman (2008) estimated that industry revenue in 2006-07 could have been of the order of AUD 1.4 billion and industry gross value added around AUD 682 million. Using a General Equilibrium (GE) modelling approach, they concluded that the economic footprint of the spatial information industry was larger.<sup>53</sup>

ACIL Tasman used a value-added approach based on General Equilibrium (GE) modelling, with input data derived from case studies, which were used as a guide to estimating the direct impacts of spatial information on selected sectors in the Australian economy, and a GE approach to modelling economy-wide impacts. A potential issue with this approach is how to scale from case studies to sector-wide impacts (i.e. understanding the relationship between the cases and the sector).

PwC (2010) computed the welfare impacts of the free provision of Geoscience Australia (GA) topographical data relative to cost recovery and noted: an agency loss of AUD 13.3 million per annum in revenue foregone; a gain for government users of AUD 10 million per annum; a gain for private consumers of AUD 8 million per annum; and an overall increase in net welfare of AUD 4.7 million per annum. Extrapolating this to annual Australian government expenditure on fundamental spatial data of around AUD 70 million would suggest net welfare benefits of around AUD 25 million per annum circa 2011.<sup>54</sup>

To support a review of Geoscience Australia (GA), the Department of Finance and Deregulation commissioned ACIL Tasman to report on the economic value of the core areas of GA's work, including that relating to pre-competitive geological information on petroleum and minerals, and its work in gathering, processing and disseminating geospatial, earth monitoring and groundwater information.<sup>55</sup>

Overall, ACIL Tasman considered that a plausible estimate of the non-productivity benefits of geospatial, earth monitoring, groundwater and hazards information is AUD 1.7 billion per annum. They did not attempt to assign a particular percentage of this benefit to GA, given the difficulty in attributing value between various fundamental data acquisition and value adding activities. However, they claimed that a large portion was likely to be attributable to GA.

---

53 ACIL Tasman (2008) *The Value of Spatial Information*, Spatial Information Systems Limited. Available [http://www.crcsi.com.au/Documents/ACILTasmanReport\\_full.aspx](http://www.crcsi.com.au/Documents/ACILTasmanReport_full.aspx)

54 Houghton, J.W. (2011) *Costs and Benefits of Data Provision*, Report to the Australian National Data Services, Canberra. Available <http://www.ands.org.au/resource/cost-benefit.html>.

55 Department of Finance and Deregulation (2011) *Strategic Review of Geoscience Australia*, Canberra.

### National statistics, hydrological and spatial data

Looking at national statistics and hydrological data, as well as spatial data, Houghton (2011), explored the economic impacts of three Australian government agencies making government data available for free and using open licensing (Creative Commons CC-BY).<sup>56</sup> The study focused on three elements:

- The costs and cost savings experienced by PSI producing agencies involved in the provision of free and open access to information;
- The costs and cost savings experienced by the users of PSI in accessing, using and re-using the information; and
- The potential wider economic and social impacts of freely accessible PSI, in the form of consumer welfare and the impacts of increased access and use, as measured by increased downloads, on returns to expenditure on data production.

The case studies presented demonstrated that even the subset of benefits that can be measured outweigh the costs of making PSI more freely and openly available. It is also clear that it is not simply about access prices, but also about the transaction costs involved. Standardised and unrestrictive licensing, such as Creative Commons, and data standards are crucial in enabling access that is truly open (i.e. free, immediate and unrestricted).

For example, Houghton (2011), found that the net cost to the Australian Bureau of Statistics (ABS) of making publications and statistics freely available online and adopting Creative Commons licensing was likely to have been around AUD 3.5 million per annum at 2005-06 prices and levels of activity, but the immediate cost savings for users were likely to have been around AUD 5 million per annum. The wider impacts in terms of additional use and uses bring substantial additional returns, with estimates suggesting overall costs associated with free online access to ABS publications and data online and unrestrictive standard licensing of around AUD 4.6 million per annum and measurable annualised benefits of perhaps AUD 25 million (i.e. more than five times the costs).

While data were more limited, there appears to have been an even more compelling case for making fundamental geospatial data freely available. Approximate estimates suggested possible agency savings on transactions of around AUD 375,000 per annum and revenue losses of AUD 1.3 million per annum. Conversely, users' savings include both (approximately AUD 1.7 million per annum). The impacts of the observed increases in use, as indicated by trends in scheduled datasets available and delivered over the period 2001-02 through 2005-06, on average social returns to annual expenditure on data collection suggest an increase in social returns of some AUD 15 million. Hence, the benefits may have been around 13 times the costs in terms of revenue foregone.

Of course, the relative cost-benefits apply to the form of PSI involved and do not reflect in any way on the performance of the producing agencies. Some forms of PSI underpin major industries and contribute to their growth and prosperity. Other forms of PSI may have an important influence on policy decisions, but the economic impacts may be more limited and difficult to trace.

---

<sup>56</sup> Houghton, J.W. (2011) *Costs and Benefits of Data Provision*, Report to the Australian National Data Services, Canberra. Available <http://www.ands.org.au/resource/cost-benefit.html>.

### 3.2.3 Open science data

Major examples of the enormous advantages of open science data, such as the Human Genome Project, are often cited, and much is made of the importance of open science data for the transparency, accountability and efficiency of scientific research. The benefits of open science data for collaboration between public and private sector researchers, and for the transfer of knowledge from public to private sectors is often discussed, but few have attempted to put an economic value on open science data.<sup>57</sup>

A partial exception is a recent series of studies exploring the value and impact of curating and sharing research data through three well-established UK research data centres – the Archaeological Data Service, the Economic and Social Data Services, and the British Atmospheric Data Centre. In summarising the results, Beagrie and Houghton (2014) noted that their economic analysis indicated that:

- Very significant increases in research, teaching and studying efficiency were realised by the users as a result of their use of the data centres;
- The value to users exceeds the investment made in data sharing and curation via the centres in all three cases; and
- By facilitating additional use, the data centres significantly increase the measurable returns on investment in the creation/collection of the data hosted.<sup>58</sup>

These studies show that the direct benefits to the data producers and users exceed the costs of curation and sharing, but do not estimate the very substantial wider benefits arising from the uses underpinned by the data (e.g. the use of atmospheric data in climate change modelling and preparedness).

### 3.2.4 What these studies show

Economic structures vary from country to country, making direct conversions problematic. Nevertheless, assuming similar levels of government data investment and use in Australia as those in the EU, based simply on GDP shares, it is possible to get a sense of the range of market and wider values of government data in Australia from these many and varied studies.

The more conservative bottom-up DotEcon (2006) estimates would suggest a PSI net value of the order of AUD 625 million to AUD 1.2 billion per annum in Australia in today's prices, while the Deloitte/BIS (2013) estimates would suggest a PSI direct value of around AUD 1.9 billion and a total value of up to AUD 7 billion per annum in Australia in today's prices.<sup>60</sup>

---

57 There is, however, an extensive related literature on the costs and benefits of open access scientific publishing, some of which has involved attempts to measure the value of enhanced access to research findings (e.g. Houghton, J.W. and Sheehan, P. (2009) 'Estimating the potential impacts of open access to research findings,' *Economic Analysis and Policy* 39(1). Available [http://www.eap-journal.com/vol\\_39\\_iss\\_1.php](http://www.eap-journal.com/vol_39_iss_1.php)).

58 Beagrie, N. and Houghton, J.W. (2014) *The Value and Impact of Data Sharing and Curation: A synthesis of three recent studies of UK research data centres*, Jisc, Bristol and London. Available [http://repository.jisc.ac.uk/5568/1/IDF308\\_-\\_Digital\\_Infrastructure\\_Directions\\_Report%2C\\_Jan14\\_v1-04.pdf](http://repository.jisc.ac.uk/5568/1/IDF308_-_Digital_Infrastructure_Directions_Report%2C_Jan14_v1-04.pdf).

59 DotEcon (2006) *The commercial use of public information (CUPI)*, Report offt861, Office of Fair Trading, London. Available <http://www.opsi.gov.uk/advice/poi/oft-cupi.pdf>.

60 Deloitte (2013) *Market Assessment of Public Sector Information*, Report to the Department for Business, Innovation and Skills, London. Available [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf)

The larger top-down PIRA (2000) estimates would be equivalent to an investment value in PSI of the order of AUD 3 billion and a use value of around AUD 22 billion per annum in Australia in today's prices,<sup>61</sup> and the 2006 MEPSIR study estimates of the PSI market value would be equivalent to around AUD 3.9 billion per annum in Australia in today's prices.<sup>62</sup>

**Table 3 Summary of international study findings (per annum)**

<i>Value measured</i>	<i>AUD per annum circa 2013</i>
Net value	\$625 million to \$1.2 billion
Investment value	\$3 billion
Market value	\$3.9 billion to \$4.5 billion
Direct value	\$1.9 billion to \$7 billion
Use value	\$22 billion
Aggregate direct and indirect value	\$25 billion

Note: Study results scaled to Australia based on pro rata GDP shares.

Source: Various sources. Authors' analysis.

Noting rapid growth in PSI markets and scaling these older estimates to 2010, Vickery's (2010) European estimates would suggest a PSI market value of the order of AUD 4.2 billion in Australia in today's prices, with aggregate direct and indirect economic impacts of around AUD 19 billion and possibly up to AUD 25 billion.<sup>63</sup> With continued growth since 2010, today one might expect this to translate to a market of perhaps AUD 4.5 billion or more and aggregate direct and indirect impacts of some AUD 25 billion or more per annum in Australia from government data alone.

### 3.3 Estimates of the value of open data

In this section, we explore two approaches to exploring the value of the open data opportunity (i.e. both the overall value and the incremental value yet to be realised).

The first involves estimates of returns to investment in the creation/collection of data and the impacts on returns of increased data accessibility. Adopting a very conservative economic approach and focusing on government and publicly funded research data only, this provides what is very much a lower bound estimate of potential value and impacts.

The second involves extrapolation of national and sector impact estimates from a study by the McKinsey Global Institute, which includes the entire range of potentially open data. Beginning our investigations thinking that this was likely an overestimate, we came to believe that, providing one uses it to derive estimates of the additional gain from additional measures on open data, it was a reasonable, and possibly even conservative estimate. We have already seen how the gains to open data may exhibit strong and unanticipated non-linearities (when traditional estimation methods – including our own – assume linearity). Further, our own case studies suggest large gains from more open data, which would be intensified by international coordination of the same.

61 PIRA (2000) *Commercial exploitation of Europe's public sector information*, European Commission, Brussels. Available [ftp://ftp.cordis.europa.eu/pub/econtent/docs/2000\\_1558\\_en.pdf](ftp://ftp.cordis.europa.eu/pub/econtent/docs/2000_1558_en.pdf)

62 Dekkers, M., Polman, F., te Velde, R. and de Vries, M. (2006) *MEPSIR: Measuring European Public Sector Information Resources*, European Commission, Brussels. Available [http://ec.europa.eu/information\\_society/newsroom/cf/document.cfm?doc\\_id=1197](http://ec.europa.eu/information_society/newsroom/cf/document.cfm?doc_id=1197).

63 Vickery, G. (2010) *Review of recent studies on PSI re-use and related market developments*, European Commission, Brussels. Available [http://ec.europa.eu/information\\_society/newsroom/cf/document.cfm?doc\\_id=1093](http://ec.europa.eu/information_society/newsroom/cf/document.cfm?doc_id=1093)

### 3.3.1 Increasing the returns to investment in data

Exploring the benefits of open access scientific publishing, Houghton et al. (2009), developed and applied a modified Solow-Swan model to estimate the potential impacts of more open access to research publications and data on social returns to R&D.<sup>64</sup> The standard Solow-Swan model makes some key simplifying assumptions, including the assumption that all knowledge is equally accessible to all entities that could make productive use of it. Obviously, this is not realistic. In the real world, there are limits and barriers to access. So they introduced *accessibility* into the standard model as a negative or friction variable, and then looked at the impact on returns to R&D of reducing the friction by increasing accessibility (for details of the method see Houghton and Sheehan 2009).<sup>65</sup>

Using such an approach to estimate the increase in returns to expenditure on data collection/creation resulting from increases in accessibility and use of that data, depends on a number of key data inputs. Crucially, these include:

- Annual expenditure on data collection/creation and preparation for use for both government data and research (science) data, and the rates of growth of those expenditures (e.g. as indicated by expenditure trends);
- The percentage increase in accessibility and use arising from making data open (e.g. as indicated by data access and downloads);
- The average (and if possible marginal) rate of return on investment in data creation/collection; and
- The useful life (shelf-life) and rate of depreciation of the data.

It is also necessary to make a series of judgements about discount rates.

While collecting such information about the many and various types and uses of data would be a daunting, perhaps impossible task, it is possible to derive some estimates of the main parameters sufficient to enable the development of range estimates of the value of open government and research data. On the basis of preliminary analysis and triangulation, we adopt the following estimates.

#### Government data creation/collection expenditures

It is difficult to estimate expenditure on data creation/collection as it is rarely individually identified, often being an integral part of operational activities rather than a defined activity in its own right. There are also some costs associated with preparing the data for use and making it available, which are also buried in more aggregated activity costs. Nevertheless, there are partial sources from which to draw. Here, by way of triangulation, we focus on three.

*First, previous analysis and experience:* Previous work with the Australian Bureau of Statistics (ABS) suggested that data creation/collection accounts for around 50% of the agency's government appropriations (Houghton 2011).<sup>66</sup> Scaled to 2013, that would be equivalent to around AUD 195 million per annum.

---

64 Houghton, J.W. and Sheehan, P. (2009) 'Estimating the potential impacts of open access to research findings,' *Economic Analysis and Policy* 39(1). Available [http://www.eap-journal.com/vol\\_39\\_iss\\_1.php](http://www.eap-journal.com/vol_39_iss_1.php); and Houghton, J.W., Rasmussen, B., Sheehan, P.J., Oppenheim, C., Morris, A., Creaser, C., Greenwood, H., Summers, M. and Gourlay, A. (2009) *Economic Implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits*, Report to The Joint Information Systems Committee (JISC) by Victoria University & Loughborough University. Available <http://www.jisc.ac.uk/publications/reports/2009/economicpublishingmodelsfinalreport.aspx>.

65 Houghton, J.W. and Sheehan, P. (2009) 'Estimating the potential impacts of open access to research findings,' *Economic Analysis and Policy* 39(1). Available [http://www.eap-journal.com/vol\\_39\\_iss\\_1.php](http://www.eap-journal.com/vol_39_iss_1.php).

66 Houghton, J.W. (2011) *Costs and Benefits of Data Provision*, Report to the Australian National Data Services, Canberra. Available <http://www.ands.org.au/resource/cost-benefit.html>.



It has been reported that Geoscience Australia (GA) spent around AUD 70 million per annum on the creation/ collection of scheduled fundamental spatial data circa 2005-06, and a further AUD 20 million was spent by other agencies (SPDE 2007).<sup>67</sup> Scaled to 2013, that would be equivalent to around AUD 115 million per annum. By comparison, GA's annual government appropriation was around AUD 118 million in the 2012-13 budget.

If we assume the Bureau of Meteorology spends a similar share of government appropriations on data creation/collection as does the ABS, then scaled to 2013 that would be equivalent to around AUD 175 million per annum.

There are other significant Commonwealth Government departments (e.g. the Department of Agriculture, Forestry and Fisheries; the former Department of Sustainability, Environment, Water, Population and Communities; the Department of Resources, Energy and Tourism; the Department of Health and Aging; etc.), which for the purposes of preliminary estimation we assume spend 15-20% of their government appropriations on data creation/collection. Scaled to 2013, that would be equivalent to around AUD 2 billion per annum. If we assume that these agencies account for 50% of Commonwealth Government data creation/collection expenditure, then in total, Commonwealth Government data spending might be around AUD 5 billion per annum.

*Second, previous international analysis:* A number of the international studies reviewed above provide estimates of the investment value of government data. For example, PIRA (2000), estimated a PSI investment value of EUR 9.5 billion in the EU countries in 1999. While economic structures vary, on a simple pro rata basis, scaled to GDP, that would be equivalent to a little over AUD 3 billion in 2013 prices – around 4% of Commonwealth Government Budget Portfolio Appropriations for 2013-14.

*Third, Commonwealth Government Budget Appropriations:* Simply taking Commonwealth Government Budget Appropriations for 2013-14 and allocating shares to departments, assuming that 15% of government appropriations was spent on data creation/collection by the four major data generating portfolios, and 5% of government appropriations was spent on data creation/collection by the others, would be equivalent to around AUD 5 billion per annum in 2013 prices (6.6% of budget appropriations).

Splitting the difference between the second and third approaches, an across the board 5% of appropriations would be equivalent to around AUD 4 billion per annum. By comparison, government expenditure on ICT is around AUD 5 billion per annum.<sup>68</sup>

Across these three methods, the estimates for Commonwealth Government data creation/collection expenditure converge on some AUD 4 billion to AUD 5 billion per annum. Assuming the State, Territory and Local Governments might spend as much again suggests total government data creation/collection expenditure of perhaps AUD 8 billion to AUD 10 billion per annum.

### **Research data expenditures**

Expenditure on R&D is partly private and partly public, and we seek to include data arising from publicly funded research in our estimates of the impacts of open research data. The sum of government and higher education spending on R&D (by sector of execution) is currently around AUD 12 billion per annum.

In a series of studies of research data centres spanning the humanities, social, and natural sciences, Beagrie and Houghton (2012, 2013a, 2013b, 2014) found that working with data (i.e. "creating, manipulating and

---

67 SPDE (2007) *Spatial Data Policy Executive Annual Report 2005-2006*, Office of Spatial Data Management, Canberra.

68 [http://www.finance.gov.au/sites/default/files/ict\\_expenditure\\_report\\_2008-09\\_-\\_2010-11.pdf](http://www.finance.gov.au/sites/default/files/ict_expenditure_report_2008-09_-_2010-11.pdf)

analysing data") accounted for 41-76% of reported research time.<sup>69</sup> So, for the purposes of preliminary estimates, we assume that 20% of research expenditure relates to research data creation/collection (i.e. one-third of the reported mid-range data time). Scaled to 2013, that would be equivalent to around AUD 2.5 billion per annum.

Hence the sum of government and research data creation/collection expenditure in Australia might be of the order of AUD 10-12 billion per annum – giving a mid range estimate of around AUD 11 billion per annum.

### Returns to investment in data

Returns to R&D vary widely, but a characteristic finding is that returns are high – often in the region of 20-60% (Bernstein and Nadiri 1991; Griliches 1995; Industry Commission 1995; Salter and Martin 2001; Scott et al. 2002; Dowrick 2003; Shanks and Zheng 2006; Martin and Tang 2007; Sveikauskas 2007; Hall et al. 2009).<sup>70</sup>

To the extent that government-funded PSI producing agencies' activities can be seen as similar to research (e.g. producing survey-based or observational data), their outputs may be supposed to exhibit similar returns. Hence, it is possible to adopt a similar approach to estimating the value of open government and research data. For the purposes of preliminary analysis, we explore the lower bound 20% and upper bound 60% average returns.

We proceed on the basis of these data, estimating the increase in annual return on investment in the data creation/collection due to additional use. As these returns are recurring during the useful life of the data, we use a simple Perpetual Inventory Method to estimate the overall value of the impacts.

---

69 Beagrie, N., Houghton, J.W., Palaiologk, A. and Williams, P. (2012) *Economic Evaluation of Research Data Service Infrastructure: A Study for The ESRC*, Economic And Social Research Council. Available: <http://esrc.ac.uk/research/evaluation-impact/impact-evaluation/economic-impact-evaluation.aspx>; Beagrie, N. and Houghton J.W. (2013a) *The Value and Impact of the Archaeology Data Service: a study and methods for enhancing sustainability*. Available: <http://jisc.ac.uk/whatwedo/programmes/preservation/ADSImpact.aspx>; Beagrie, N. and Houghton J.W. (2013b) *The Value and Impact of the British Atmospheric Data Centre*, Jisc. Available: [http://jisc.ac.uk/whatwedo/programmes/di\\_directions/strategicdirections/badc.aspx](http://jisc.ac.uk/whatwedo/programmes/di_directions/strategicdirections/badc.aspx); and Beagrie, N. and Houghton J.W. (2014) *The Value and Impact of Data Sharing and Curation: A synthesis of three recent studies of UK research data centres*, Jisc. Available <http://repository.jisc.ac.uk/5568/>

70 Useful reviews of this literature include Griliches, Z. (1995) 'R&D and productivity: Econometric Results and Measurement Issues,' In Stoneman, P. (Ed.) (1995) *Handbook of The Economics of Innovation and Technological Change*. Blackwell, Oxford, pp52–89.; Salter, A.J. and Martin, B.R. (2001) 'The economic benefits of publicly funded basic research: a critical review,' *Research Policy* 30(3), pp509-532.; Scott, A., Steyn, G., Geuna, A., Brusoni, S. and Steinmueller, E. (2002) *The Economic Returns to Basic Research and the Benefits of University-Industry Relationships*. Report to the Office of Science and Technology, London.; Shanks, S. and Zheng, S. (2006) *Econometric modeling of R&D and Australia's productivity*, Staff Working Paper, Productivity Commission, Canberra.; Martin, B.R. and Tang, P. (2007) *The benefits of publicly funded research*, SWEPS Paper No. 161, Science Policy Research Unit, University of Sussex, Brighton.; Sveikauskas, L. (2007) *R&D and Productivity Growth: A Review of the Literature*, US Bureau of Labor Statistics Working Paper 408, BLS, Washington, DC.; Hall, B.H., Mairesse, J. and Mohnen, P. (2009) *Measuring the returns to R&D*, NBER Working Paper 15622, NBER, Cambridge MA.; Hall, B.H., Mairesse, J. and Mohnen, P. (2010) *Measuring the returns to R&D*, in eds. Hall, B.H. and Rosenburg, N. (2010) *Handbook of the Economics of Innovation*, North Holland.

Drawing on preliminary work on the UK R&D Satellite Account (Evans et al. 2008),<sup>71</sup> and following the lead of the US R&D Satellite Account (Sveikauskas 2007),<sup>72</sup> we depreciate publicly-funded government and research data at 10% per annum, and we set the useful life of the data created each year at an average of 10 years – although, of course, the useful life of data can be much shorter and/or much longer depending on data type and use. Following Mansfield (1991, 1998),<sup>73</sup> we distribute the returns normally over five years from year 1 through year 5. Applying a 4% discount rate to estimate net present value, we then model the recurring increase in returns to expenditure on the data due to increases in accessibility arising from making the data more open.

**Table 4 Parameters used for modelling**

<i>Item</i>	<i>Value</i>
Expenditure on government data (AUD millions per annum)	8,919
Expenditure on research data (AUD millions per annum)	2,450
Growth in data creation/collection expenditure (per cent per annum)	3%
Returns to data creation/collection expenditure (per cent per annum)	20% to 60%
Increase in accessibility and use of the data from more open data (per cent)	50% to 100%
Useful life of the data (years, averaged across all data types)	10
Rate of depreciation of the underlying stock of data (per cent per annum)	10%
Discount rate to estimate Net Present Value (per cent)	4%

*Source: Various sources. Authors' analysis.*

As the change to open data is a one-off (step-change) increase in data accessibility and use, it creates growth rate effects. Hence these estimates can provide insight into the contribution of more open data to economic growth, as well as providing the basis for estimates of both existing returns and the potential additional returns from making government and research data more open.

### Modelling scenarios and findings

Under the modelling assumptions outlined in Table 4 (above), the implied returns from current investment in Australian government and publicly funded research data would amount to around AUD 240 billion over 20 years in Net Present Value (NPV), ranging from AUD 120 billion at the lower bound 20% return to AUD 360 billion at the upper bound 60% return.<sup>74</sup> At the mid-point, the return from one year's data spending would be around AUD 17 billion NPV over 20 years, a return worth 1.5 times the investment.

Exploring the additional returns that might arise from increased data accessibility and use, we find that a 50% increase in access and use might be worth around AUD 120 billion over 20 years (NPV), with a lower bound of AUD 60 billion and an upper bound of AUD 180 billion.

71 Evans, P., Hatcher, M. and Whittard, D. (2008) 'The preliminary satellite account for the UK: a sensitivity analysis', *Economic & Labour Market Review*, 2(9), September 2008, 37-43.

72 Sveikauskas, L. (2007) *R&D and Productivity Growth: A Review of the Literature*, US Bureau of Labor Statistics, Washington DC., Working Paper 408.

73 Mansfield, E. (1991) 'Academic research and industrial innovation,' *Research Policy* 20(1), 1991, pp.1-12; and Mansfield, E. (1998) 'Academic research and industrial innovation: an update of empirical findings,' *Research Policy* 26(7/8), 1998, pp.773-776.

74 A 20-year period is used to capture the value of data that has a multi-year 'shelf-life', with returns to investment accruing over a number of years.

Under these assumptions, a doubling of accessibility and use would see the returns to investment in Australian government and research data rise by around AUD 240 billion over 20 years (NPV), ranging from an increase of AUD 120 billion at the lower bound 20% return to an increase of AUD 360 billion at the upper bound 60% return. These estimates are no more than indicative. There is no reason to suppose that doubling accessibility and use is the maximum achievable increase. Use could increase by a factor of three, five or ten. Moreover, these estimates are probably conservative (Box 6).

The G20 governments have set the target of increasing growth by 2% over the next five years, saying:

*We will develop ambitious but realistic policies with the aim to lift our collective GDP by more than 2% above the trajectory implied by current policies over the coming five years.*<sup>75</sup>

Under our conservative modelling assumptions and focusing on government and research data alone, a doubling of accessibility and use would see the returns to investment in Australian government and research data rise by around AUD 19 billion over five years (NPV), ranging from an increase of AUD 10 billion at the lower bound 20% return to an increase of AUD 29 billion at the upper bound 60% return.

### **Box 6 Conservative modelling assumptions**

There are a number of reasons why the assumptions underlying these estimates are likely to underestimate the value and impact of open data.

*Constant returns:* While there are many different types of data and different uses of that data, and it is likely that some will generate increasing and some decreasing returns, it is widely held that information is often characterised by increasing returns. In medicine, for example, something like Metcalfe's law of networks may apply – in which data grows not just incrementally, but sometimes by leaps and bounds in its usefulness as it is linked to other data. Thus, for instance, genotype data is not much use until linked to phenotype data, and where this can be linked to unit patient records, phenotypical data then gains much greater depth. Link this to Pharmaceutical Benefits Scheme (PBS) data and you have a powerful engine for driving productivity breakthroughs in many areas as patterns emerge reflecting drug efficacy, safety and drug interactions (See also Section 3.1).

*Limited to government and publicly funded research data:* There is much data that is not included in the estimates, which are limited to government and publicly funded research data (See Figure 5). Indeed, it may well be that most data and, perhaps, the most valuable/useful data arise from privately funded R&D and in the course of business, NGO and non-profit sector activities.

*Source: Lateral Economics analysis.*

Such a doubling of accessibility and use would be equivalent to an additional 0.27% of Australia's cumulative GDP over the next five years, ranging from 0.14% to 0.41% of GDP. On these conservative estimates of increased returns to investment in data, doubling the accessibility and use of Australian government data and publicly funded research data might contribute around one-seventh of the G20's 2% additional growth target.

### **3.3.2 Valuing the impacts of open data**

Broadening the net to include all forms of data, public and private, McKinsey Global Institute (2013), identified the potential economic value that could follow from enhancing the accessibility of data. The study suggested that:

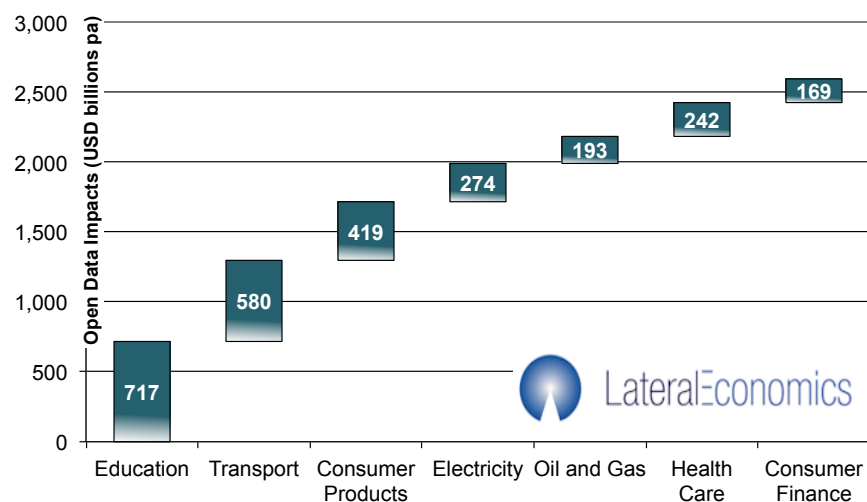
---

<sup>75</sup> Communiqué, Meeting of Finance Ministers and Central Bank Governors, Sydney, 22-23 February 2014.

“...making data more widely available in shareable formats, and more ‘liquid’ (open, widely available, and in shareable formats) has the potential to unlock large amounts of economic value, by improving the efficiency and effectiveness of existing processes; making possible new products, services, and markets; and creating value for individual consumers and citizens”.<sup>76</sup>

The monetised values in the report are expressed in terms of annual economic surplus, in 2013 US dollars, and are not the discounted value of future cash flows. The report does not attempt to estimate open data’s relative contribution to the overall impacts facilitated. Rather, the estimates represent the total value created. The study suggested that open data has the potential to unlock approximately USD 3.2 trillion in additional value annually across the seven domains of the economy that the study examined. In McKinsey’s view these benefits could be recurring because “individuals perceive the benefits of open data, they will help to improve the accuracy and detail of information available, thus increasing the value of the data and the benefits.”<sup>77</sup>

**Figure 8 Value of open data globally (USD billions pa)**



Source: Based on McKinsey Global Institute (2013) *Open data: Unlocking innovation and performance with liquid information*. New York.

To more fully comprehend the *incremental* benefits that could arise from open data, we explore the changes in world output under two scenarios: a business as usual scenario, and one in which an open data regime is implemented and the impacts realised.

Under a business as usual scenario we assume that growth in world gross domestic product (GDP) continues at 2% per annum – based on World Bank data historical trends. Based on the McKinsey estimate of the value of open data of USD 3.2 trillion per annum, the worldwide adoption of open data could increase output by some USD 16 trillion cumulatively over five years in current prices.

Given that the impact identified by McKinsey is output rather than value added and they do not attempt to attribute a share to open data directly, we must scale these impacts to estimate their contribution to GDP.

<sup>76</sup> McKinsey Global Institute (2013) *Open data: Unlocking innovation and performance with liquid information*. Authors’ analysis. New York.

<sup>77</sup> McKinsey Global Institute (2013) *Open data: Unlocking innovation and performance with liquid information*. Authors’ analysis. New York.

Assuming that two-thirds of the impact might be value added and one-half attributable to open data, the contribution to cumulative GDP growth might be around 1.4% over the next five years. Perhaps four-fifths of this remains to be realised, suggesting that a vigorous open data agenda might add around 1.12% to GDP over the next five years.

### **Box 7      How these impacts are realised**

Within each domain, McKinsey examines the effect of 'levers', unique to each domain, that effect benefits from the implementation of open data regimes.

*Education:* The largest potential benefit could come from improvements to instruction through expanding the most effective strategies and tools for teaching with resulting higher skills and earnings. Open data also improves choices on academic or vocational pathways.

*Transport:* Open data can increase productivity and reduce travel times for passengers and freight, through adjusting schedules to optimise operations based on industry-wide benchmarks, also reducing infrastructure investment.

*Consumer products:* Increased price transparency informs and intensifies competition. Open data could generate enhanced sales and loyalty for manufacturers and retailers, by greater differentiation of consumer types and customised store layouts.

*Electricity:* Open data can enable consumers to discover energy-efficiency opportunities. Also, by providing consumers and businesses with detailed data about their energy consumption, utilities can benefit from sharing benchmarking data to improve project management and operations.

*Oil and gas:* The potential benefits include improvements to investment decisions about areas to explore for new reserves and build downstream facilities. Sharing consumption data can help consumers make better-informed decisions about energy use, thereby contributing, for example, to possible reductions in natural gas consumption.

*Health care:* The potential sources of value in this domain arise from enabling people to take an active role in disease prevention and treatment, as well as helping providers determine treatment that is the most timely and appropriate for each patient. It could also result in the matching of patients with the most appropriate providers and identifying new therapies and approaches to delivering care.

*Consumer finance:* The use of open data in consumer finance could assist in product design, risk assessment of consumers who do not have a credit history, fraud prevention and detection, as well as assist consumers to choose among a wide range of complex financial products.

*Source: Based on McKinsey Global Institute (2013) Open data: Unlocking innovation and performance with liquid information. Authors' analysis.*

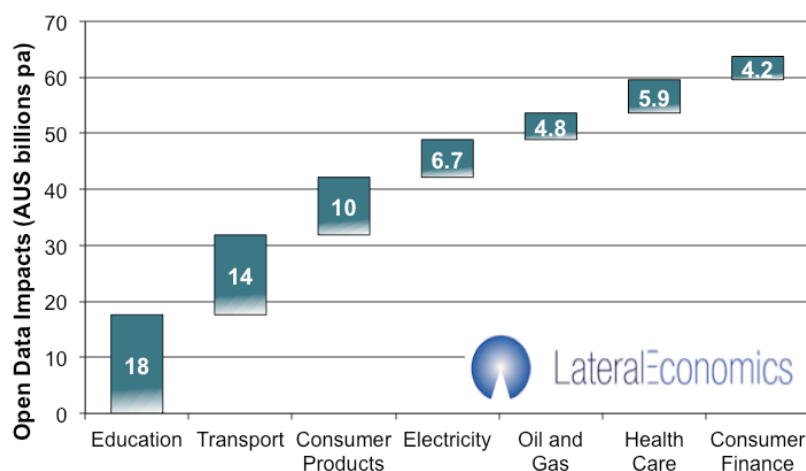
### **Potential implications of open data for Australia**

Although the McKinsey study did not provide information on Australia's share of the potential value to flow from open data, its portion can be estimated.

McKinsey suggested that around 46% of the potential value of open data would accrue to the Rest of the World Group (ROW) of countries – a group excluding the US and Europe. Based on its GDP share of this ROW group of countries, around 4.3% of this value can be attributed to Australia. On this basis, we estimate the potential value of open data to Australia at almost AUD 64 billion per annum.<sup>78</sup>

Using the same approach as above to illustrate the effects of open data, we find that national output could have increased by around AUD 320 billion cumulatively over the next five years in 2013 prices.

**Figure 9 Value of open data for Australia (AUD billions pa)**



*Note: Sector impact shares are simply pro rata on share of GDP and take no account of structural differences between economies. Hence, they should be treated as no more than broadly indicative.*

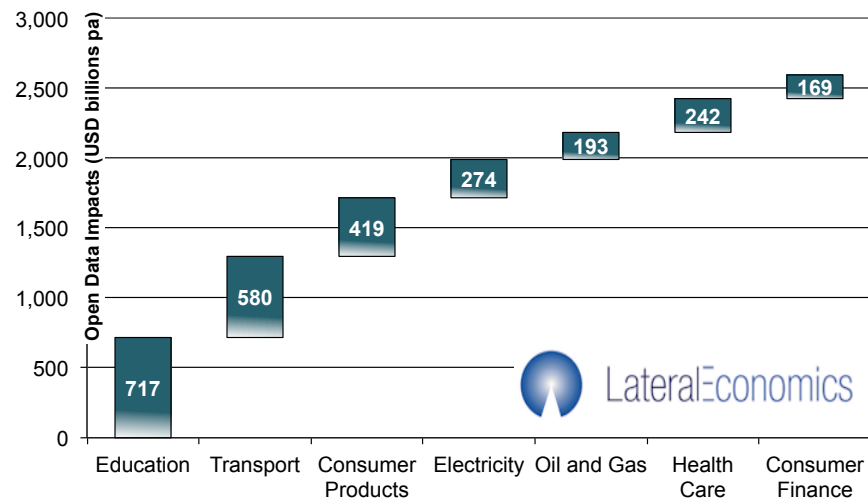
*Source: Based on McKinsey Global Institute (2013) Open data: Unlocking innovation and performance with liquid information. New York.*

Again assuming that two-thirds of the impact might be value added and one-half attributable to open data, the contribution of open data (both existing and potential) to Australia's cumulative GDP growth might be around 1.3% over the next five years. With perhaps three-quarters of this remaining to be realised, a vigorous open data agenda might increase GDP by around 1% over the next five years.

### Potential implications of open data for G20 countries

Based on the G20's GDP share, around 80% of the total value of open data can be attributed to the G20 countries. Hence, the potential value of open data to the G20 might be almost USD 2.6 trillion per annum. Using the same approach as above to illustrate the effects of open data, we find that G20 output could have increased by around USD 13 trillion cumulatively over the next five years in 2013 prices.

<sup>78</sup> These numbers have been calculated from McKinsey figures which are given in US dollars. However the US and Australian dollar were around parity at the time of the estimation.

**Figure 10 Value of open data for the G20 (USD billions pa)**

Source: Based on McKinsey Global Institute estimates, pro rata on GDP shares. Authors' analysis.

The G20 countries have set the target of lifting aggregate GDP by 2% above the trajectory implied by current policies over the coming five years. Again assuming that two-thirds of the impact might be value added and one-half attributable to open data, the contribution of open data to aggregate G20 countries' cumulative GDP might be around 1.3% over the next five years, of which perhaps 1.1% remains to be realised.

Thus open data could contribute almost 55% of the G20 growth target of 2%.

### 3.4 Summarising these findings

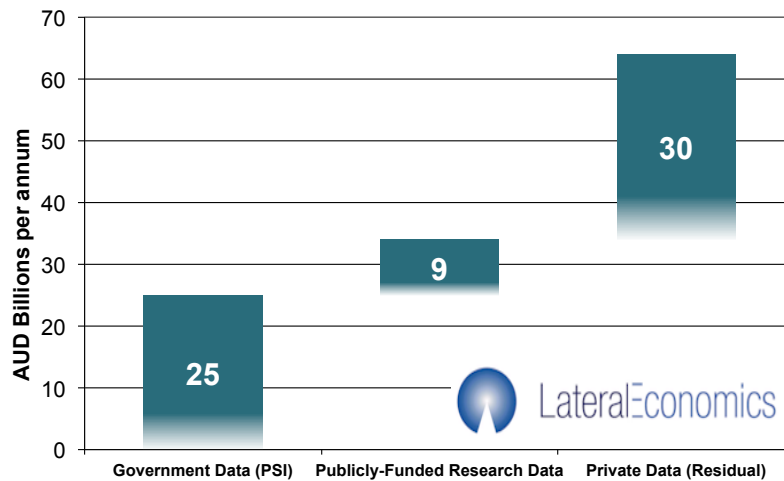
While the various estimates use different approaches and measure different forms of value, they do create a coherent overall picture.

- Based on the previous studies analysed in section 3.1, we estimate the aggregate direct and indirect value of government data in Australia at up to AUD 25 billion per annum;
- Based on our analysis of the returns to investment in government and research data, assuming a doubling of accessibility and use, we estimate the return on Australia's investment in government and research data at around AUD 34 billion per annum; and
- Based on our extrapolation of estimates from the McKinsey Global Institute, we estimate the potential economic value of open data (i.e. including government, research, private and business data) for Australia at up to AUD 64 billion per annum (at parity).

While what is being measured varies, an approximate idea of the way these elements might combine is shown in Figure 11.



**Figure 11 Estimates of the value of open data (AUD billions pa)**



Source: Lateral Economics' analysis.

While some of this value is already being obtained, there is much remaining to be realised through adopting a more aggressive open data agenda. For example, we estimate that open data alone could contribute as much as 55% of the G20 growth target of an additional 2% growth over next five years.

### 3.5 Case studies and examples

It is impossible to account for all of the possible impacts and benefits that can arise through open data. Nevertheless, it is important to understand how the many different forms of economic and wider benefits can arise, in order to flesh out the economic estimates and further our understanding of the mechanisms through which value can be realised and the many barriers facing that realisation.

The case studies presented in the Annex demonstrate how open data can create value. These examples have been selected to coincide with the themes of the G20 Agenda (Figure 12). The case studies and examples are also used to demonstrate the different ways in which governments can act to create value through open data, both individually and collectively, through such forums as the G20.

**Figure 12 Mapping case studies to the G20 Agenda**

CASE STUDIES	Strong, Sustainable and Balanced Growth			CASE STUDIES
ANNEX 1 Trade Data International	Trade	Boosting Growth through the Private Sector	Investment and Infrastructure	ANNEX 7 Roads Water
	Development		Fiscal and Monetary Policy	ANNEX 3 Managing Economic Cycles
ANNEX 5 Windows on Workplaces, APS Jobs	Employment	Building Global Economic Resilience	Tax	
ANNEX 4 Anti-Corruption	Anti-Corruption		Reforming Global Institutions	
ANNEX 6 Energy	Energy		Financial Regulations	ANNEX 2 Sirca (Tick Data)

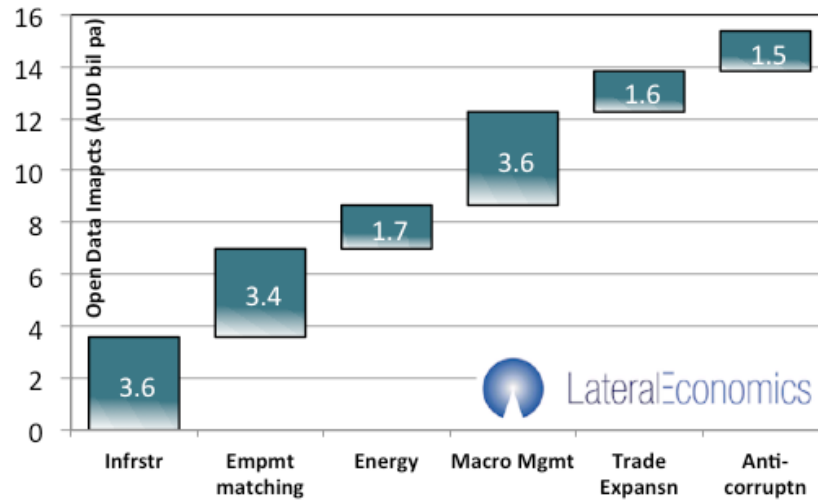
Source: *Lateral Economics Analysis*

It has not been possible to estimate the impacts of all of the open data case studies explored, but it is clear from those for which estimates have been made that their impacts are substantial. The potential captured in our case studies amounts to over AUD 15 billion per annum or around 1% of GDP which is half the G20 2% target. This is a striking result given that this is a snapshot of results rather than a comprehensive estimation of the economic value available within the G20 themes, let alone the whole economy.

We provide indicative estimates of these economic returns in the following areas:

- Trade: Open data enabling the extension of trade could add AUD 1.6 billion per annum to Australia's economy, equivalent to an increase of 0.10% in GDP (Annex 1).
- Fiscal and monetary policy: Open data enabling better management of the economic cycle could add AUD 3.6 billion per annum to Australia's economy, equivalent to an increase of 0.23% in GDP. (Annex 3)
- Anti-corruption – we estimate that open data enabling greater transparency could contribute AUD 1.5 billion per annum to Australia's economy, equivalent to an increase of 0.10% in GDP. (Annex 4)
- Employment – we estimate that open data enabling better matching of people and jobs could add AUD 3.4 billion per annum to Australia's economy, equivalent to an increase of 0.22% in GDP. (Annex 5).
- Energy – we estimate that open data could lead to improvements in investment decisions and operations management in energy supply and some areas related to demand sector, which together would add 0.11% or around AUD 1.7 billion to Australia's economy annually. (Annex 6)
- Infrastructure open data enabling better design and better use of Australia's infrastructure could contribute AUD 3.6 billion per annum to Australia's economy, equivalent to an increase of 0.23% in GDP. (Annex 7).

**Figure 13** The contribution of open data case study examples to the Australian economy (AUD billions per annum)



Source: Lateral Economics Analysis

## 4. The way forward

This section explores policy scope and directions, suggesting an open data opportunity that spans government data (PSI), data arising from publicly funded research, and private sector data. Hence the opportunity spans:

- Direct open data policies focusing on making more data openly available;
- Supporting and prosecuting policies focusing on opening publicly funded research data (e.g. through research funding agencies); and
- Establishing a regulatory and policy environment conducive to enabling the maximisation of opportunities to open private sector data.

Policies in these areas have both a national dimension (e.g. open data produced by government and publicly funded research), and an important international dimension (e.g. standardisation and harmonisation of national regimes) that can play a crucial role in fully realising the open data opportunity. Thus, there is an important role for international forums such as the G20.

### 4.1 Generic opportunities, issues and actions

The G20 countries leading the open data agenda – the UK, the US and Australia – have all implemented explicit policies to maximise the extent to which government and publicly funded research data is released and made available in a form that maximises its total value including by optimising the capacity to add value to it.

This includes such commitments as machine readable and open licensed release by default subject to appeal to an independent Information Commissioner, as well as centralised distribution and marketing of data through well advertised portals, such as [data.gov.au](http://data.gov.au).

Many other governments are also progressing to an open government agenda, nationally and through such organizations as the Open Government Partnership (OGP). Those in at the beginning in 2011 included Brazil, Indonesia, Mexico, Norway, Philippines, South Africa, the UK and US.<sup>79</sup>

While good progress on these high level 'constitutional' matters of legal enablement and entitlement have been made, there remain substantial practical obstacles to easy release of data that continue to impede progress, often very substantially.

Such things include concerns over:

- Privacy, confidentiality and other ethical concerns about open data 'in the wild';
- Security, which may extend to international obligations not to divulge information that may compromise military secrets;
- Deciding:
  - o who are the data custodians?
  - o what are their responsibilities?
  - o how should they be resourced to do open data work and how should relative priorities be determined regarding this?

---

79 See <http://www.opengovpartnership.org/>

- What data linkage should be facilitated, and what should be prevented, for instance to protect privacy or confidentiality;
- Prior copyright issues (e.g. from existing content supply contracts) and deciding on a licensing regime; and
- How agencies attend to the needs of others.

A good example of the ways in which such concerns can slow down the pace of release, of the 'data value at risk', so to speak, is provided in Box 8.

### **Box 8      Obstacles to innovation using Public Service Gazette data**

As outlined above, information in the Australian Public Service (APS) Gazette has a substantial value to which Australian firms such as [www.APSJobs.info](http://www.APSJobs.info) have added in recent times. The initial obstacle to value adding was that the Australian Public Service Commission (APSC) did not make the data available in a machine-readable format as recommend in the Government 2.0 Taskforce Report (2009) and the Australian Government's response to it.

APSJobs.info's business model was predicated on its development of successful technical methods to 'scrape' the data from pseudo-print PDFs. However, frequent changes to the formatting and layout of these files meant continuous re-development of the PDF conversion software to continue to access and add value to the data. The skills required to perform such work (data-mining and text-analytics) are in great demand, and the cost of frequently using such resources exceeded the benefits to Pivotal Analytics.

This illustrates the ease with which policies agreed at the highest levels within government can be undermined by other priorities of specific agencies and whose cooperation with both the letter and the spirit of such policies is necessary to successful implementation of the policy. It also illustrates the need for those who may contemplate adding value to the data to feel they can rely on government agencies cooperating with the efforts of outsiders seeking to add value to government data.

Indeed the greater accessibility of the data has become a concern to the Australian Public Service Commission which is currently conducting a review of the publication of the Gazette. Australia's Parliamentary Joint Committee on Human Rights fears that publishing details of dismissals through Gazetel might be violating the International Covenant on Civil and Political Rights and the Convention on the Rights of Persons with Disabilities. The head of the Australian Public Service Commission, Mr. Stephen Sedgwick noted that "[t]oday, termination information published online in the Gazette is available years after the fact, and is discoverable via a search engine".<sup>80</sup>

However, this enhanced discoverability produced by the internet is a change in degree, not in kind. As the Public Service Commissioner concedes, in the past it could always have been "tracked down in a library or archive by a motivated researcher". Yet, on the strength of the greater ease with which this can be done, we are considering overturning a century-old tenet of public sector transparency. Further, if transparency is to be curtailed to address the privacy concerns identified above, one would hope such curtailment would occur to the minimum possible extent – by maintaining the publication of the data subject to excisions of offending material – for instance, certain kinds of terminations. (See also Annex 5)

80 <http://www.smh.com.au/national/public-service/aps-gazette-employment-pages-in-doubt-20140407-367xf.html>

## 4.2 Beyond Government 2.0 and freedom of information

When it began in the US in the mid 1960s and in Australia in the early 1980s, the principal motivating force behind Freedom of Information (Fol) legislation was the civil and constitutional rights of citizens. The passage of the successor Freedom of Information Act in Australia in 2010 marks the transition to Fol being seen not just as one of civil rights and administrative hygiene, but also as one of economic efficiency. As the new 2010 Act specifies, Parliament's intention in passing it is to "increase recognition that information held by the government is to be managed for public purposes, and is a national resource".<sup>81</sup> Here as elsewhere, much of the traditional open data agenda for governments has focused on releasing existing government data to create what we have called 'public goods of opportunity'. The data exists and so its value to the society and economy will be maximised by releasing it in an open format.

A narrow interpretation of this agenda focuses on the agenda set out in the preceding section – maximising the open release of data that governments have collected or funded. Yet it is possible to imagine an open data policy agenda well beyond these relatively passive and limited aspirations. Given the power of open innovation and its tendency to feed on itself, there are ways that governments can help foster open innovation more widely in the community, both by their own treatment of their own data and by the influence they may be able to bear on others in releasing their data. We explore some of these possibilities in the following sections.

### 4.2.1 The G20 and the emergence of information standards

Information architectures are typically built around standards which define the relationship different data points and items of information. Those standards often emerge over time either spontaneously from life itself – as in the case of spoken and written languages – or as the result of deliberate collective effort as in the case of a range of standards around which the internet is built.

Gruen has argued in various contexts that governments can establish or facilitate the evolution of standards with two powerful effects. Firstly, because information travels in standards, the emergence of a standard may be crucial to increasing the value of information because that information can then acquire comparative significance. And when making decisions about alternative choices, comparative information is often the kind of information we need.

Secondly, once a standard has emerged, the very fact that it facilitates comparisons unleashes incentives for the best performers to report against it. This strengthens the use of the standard and draws information which might have previously privately held, into the public domain. Box 9 illustrates how governments might help a standard to emerge in the area of employees' experience of workplaces which could generate large gains corresponding to the employment theme of the G20 (See also Annex 5). Box 11 illustrates how the evolution of standards could improve the market for financial advice – relating to the G20 theme of Financial Regulation.

---

81 Freedom of Information Amendment (Reform) Act 2010, Objects clause. Section 3(3). Available at <http://www.comlaw.gov.au/Details/C2012C00866>.

### Box 9 Windows on Workplaces

Firms regularly survey their employees to understand how engaged they are in their work. This information has obvious value – most particularly to those considering working for the firm. It might seem obvious that firms would not want to expose this data to the public for the embarrassment it could cause them. But that does not explain why the best firms do not publish their results. If they did it could create a dynamic that forced other firms to publish their results lest people think they were covering something up.

But the problem is that there is no standard against which all firms report. As a result, no one can really compare different results. And a standard is a public, which is to say, a collective good. So "Windows on Workplaces" would involve some leader – the Prime Minister is the obvious candidate but it could be any prominent and well intentioned figurehead – to challenge the best firms to join them in developing and reporting to a standard.

The public sector is well placed to get the ball rolling. Because most public sectors have central bodies that conduct standardised surveys of employees. And as the Australian Public Service Commission (APSC) notes (2010, p19), citing the US Merit Systems Protection Board, there is "a significant positive correlation between employee engagement scores and agency performance".

*Source: Adapted from Gruen, N. (2011) "The David Solomon Lecture: Government 2.0 a couple of years on". Available <http://clubtrop-po.com.au/2011/09/27/the-david-solomon-lecture-government-2-0-a-couple-of-years-on/>.*

In addition, governments can leverage the value of their own data and operations by bringing them within emerging standards. Often all players have an incentive to do this, as for instance, those marketing web services have to optimise their websites to Google's and its competitors' search algorithms. But where this might be seen as simple marketing – maximising an agency's digital assets' exposure to the market – sometimes conforming to a standard can further enhance the value of those assets. Thus, for instance, if data is exposed to the internet so as to maximise its compatibility with a computational internet engine like Wolfram Alpha, then not only does this help the standard coalesce around a particular configuration, but the computational abilities of Wolfram Alpha then come to augment the digital assets of the relevant agency.

### Box 10 The importance of standards in facilitating comparisons

Markets for most goods are usually pretty well informed because we can inspect goods before sale and there are plenty of repeat purchasers. But what if you need a heart bypass? Your GP will recommend a surgeon. But does he know the surgeon's success rate, or the infection rates of the hospitals to choose from?

We regulate for mandatory disclosure of information to investors and consumers to tackle this kind of information asymmetry. But such disclosure regulation typically assumes that consumers and investors are in a good position to work through all the detail that's disclosed when what they really need is a way to work out which professionals they can rely on.

But if information on who to trust is so useful, why hasn't the market provided it? To be useful, information on the quality of services must enable users of the information to compare providers. And this can't be done unless providers report to the same standard. In this context the standard is a public good, which markets will often fail to produce.

While Stiglitz and Akerlof might suggest some form of regulation, Hayek reminds us of how little governments know and so how dysfunctional regulation can be, as for instance, when Financial Services Reform forces firms to produce hundred-page financial product disclosure statements that investors despair of ever understanding before throwing them in the bin. . . .

[T]here is a middle way. Governments can use their own dominance of some professional services like health and education to force much better levels of disclosure which can then potentially drive improvements in service quality as has occurred in the UK.

And sometimes all it takes is a little leadership to nudge market forces along. An energetic and prominent leader . . . could throw out a public challenge to the leaders in a field to get together and develop a standard against which to report. The best hospitals, schools and employers should jump at the chance of demonstrating their superiority.

Source: Gruen, N. (2008) *Information key to efficiency*” *Australian Financial Review*, 17th April. Available [http://www.afr.com/p/opinion/item\\_0TNtLLLUxDbDFHAXGIRUnN](http://www.afr.com/p/opinion/item_0TNtLLLUxDbDFHAXGIRUnN)

#### 4.2.2 Using government influence to establish open data accountability regimes

Government can use its funding and other influence to insist, and failing that, to seek to persuade others to maintain data and to open it according to government mandated requirements. Thus in health and education as a condition of its funding, the Commonwealth requires detailed performance information which it has begun to display publicly, for instance, on websites such as MySchool, MyHospitals. However, in contrast to best international practice, both data sets are published according to copyright notices that explicitly prevent value-adding to their data without further copyright permission.<sup>82</sup>

Governments can also influence the opening of research data, especially that arising from publicly funded research, through reporting and impact requirements. There is scope for governments to extend such influence to the spheres of data collected by the private sector (e.g. in public health and business regulation). As we live in an increasingly globalised world, there is much scope for governments to act cooperatively through international forums like the G20, to ensure that national regimes are harmonised to the maximum possible extent.

#### 4.2.3 The significance of the information architecture

Even where one has little power over actors in some eco-system, it is often possible to influence the openness of that eco-system if one has an influence over its architecture.

The architecture of the internet and the world wide web both illustrate the extent to which the architecture of a platform can influence how open it remains long after the platform is opened up to private users who may have self-interested reasons to hoard their data and indeed to actively prevent others using it. Those who designed the internet did so with nuclear war in mind. As a result, signals are sent through the network in addressed packets that can be re-routed through the network in the event of disruptions to any given node in the network.

The global phone network and the internet are both built around the need of service providers in the system – a telco or an ISP – to sell their customers access to others on the global network. They typically agree to

82 <http://www.myschool.edu.au/ContactUs/Copyright> and <http://www.myhospitals.gov.au/copyright>. Compare this with an equivalent website in the UK focused on universities <http://unistats.direct.gov.uk/open-access-data/>.



reciprocal handling of others' incoming traffic in exchange for others' handling their outgoing traffic. But even though phone network technology is largely digital, its whole architecture reflects its "analog" history.

Just as Alexander Graham Bell built it, the phone network must make connections between phones before a call takes place. No connection, no call. A service provider seeking access to all the world's telephone connections will negotiate reciprocal access with large telcos and huge networks with monopoly or monopolistically competitive market positions.

By contrast, the maximally open physical architecture of the internet has produced a maximally competitive market in the interconnect agreements which constitute the commercial infrastructure of the net. If one service provider will not reasonably negotiate an agreement to reciprocally forward packet data, other options are always available. And, since no one's indispensable, few are tempted to negotiate unreasonably. And so, miraculously all those transaction costs between service providers negotiating reciprocal access to each others' services collapse.

As OECD researchers reported recently, 99.5% of reciprocal access agreements occur informally without written contracts. Paradoxically, as competition becomes more intense or "perfect", it becomes indistinguishable from perfect co-operation – a neat trick demonstrated in economists' models a century ago.<sup>83</sup>

This architecture has proven remarkably robust to the designs of interests with huge resources and market power to establish closed networks to rival the internet. In the early days of the world-wide web, Microsoft sought to use its enormous market power to establish a proprietary 'walled garden' version of the internet that it would control. But the openness of the internet itself meant that Microsoft's objectives could not be delivered. People preferred the open internet and did not come in sufficient numbers to make Microsoft's walled garden commercially viable.

#### **Box 11 Architecture for openness: How governments can get data flowing in mineral exploration**

As new technologies have made it possible to drill in a wider range of geological formations, reservoirs have become more complex. That is raising costs and risks – estimated ratios of prospects to explored targets can be as high as 50 to 1. The sharing of information on drilling permits and on seismic and other data across companies could reduce the number of dry holes and help optimise investments.

While the widespread sharing of seismic data is unlikely, sharing among even a few companies could produce significant new value in the oil and gas industry. Governments keen on maximising resource wealth could take the lead in structuring processes for granting permits so that grants of initial drilling licenses would require greater sharing of seismic data. Sharing data on projected costs and development timetables (through third parties) could establish benchmarks that, we estimate, would reduce per-project costs by 15-25%.

*Source: Chui, Michael, Manyika, James and Van Kuijen, Steve, 2014 "What executives should know about open data", January 2014, McKinsey.*

83 Weller, D. and B. Woodcock (2013), "Internet Traffic Exchange: Market Developments and Policy Challenges", *OECD Digital Economy Papers*, No. 207, OECD Publishing, [http://www.oecd-ilibrary.org/science-and-technology/internet-traffic-exchange\\_5k918gpt130q-en](http://www.oecd-ilibrary.org/science-and-technology/internet-traffic-exchange_5k918gpt130q-en)

#### 4.2.4 Smart disclosure

Partly because of the way in which it can help a standard develop, the manner in which government data is disclosed can influence its immediate value and the extent to which it contributes to further value adding in an open eco-system. What President Obama has called “Smart disclosure” involves the “timely release of complex information and data in standardized, machine readable formats in ways that enable consumers to make informed decisions”.<sup>84</sup> Often smart disclosure will strengthen some existing standard or encourage some new standard to emerge as in the case of the US military’s “blue button” initiative that provided military personnel with their medical records in a standardised format. Such disclosures have often led to further value adding to such data from third parties who combine it with other data or services.

#### 4.2.5 Deep public private partnerships

Public private partnerships (PPPs) exist whenever governments have dealings with private firms as suppliers or partners of one kind or another. Such partnerships are common in governments delivering open data. However, it is possible to envisage ‘deep’ PPPs in which the public and private sectors cooperate in a much more integrated way.

Gruen gives the example of the personal genomic services 23andMe which provides a simple, partial genotyping service to consumers for AUD 99. Where the service has acquired around 600,000 customers from a target market which would be more than one billion, Gruen has proposed that the service be funded by a national health insurer like Medicare, that the genotypic information gathered would have a capital value worth at least the cost of acquisition to the health system and that the health system also ‘nudge’ people to opt in.<sup>85</sup>

### 4.3 Open research data

Research is producing larger and more complex data than ever before. It is imperative that these data outputs are effectively managed and shared. Better data – better described, more connected, more integrated and organized, more accessible, more easily used for new purposes – allows new questions to be investigated, larger issues to be investigated, and data landscapes to be explored.<sup>86</sup>

Governments can play a key role in making research data more openly available, especially data arising from publicly funded research. For example, government research funding can carry with it an explicit expectation that the data arising from funded research be made openly available by default – subject to specific exceptions relating to privacy, security, etc. This can be achieved through research funding councils’ mandating open data, and requiring a data plan to be submitted with the proposal (e.g. as already happens through the Australian Research Council).

There must also be practical support for open research data management and curation. An example of this is the Australian National Data Service (ANDS), which is supporting Australia’s research data environment so as to make Australian research data collections more valuable by managing, connecting, enabling discovery and supporting the reuse of this data, to enable richer research, more accountable research; more efficient use of research data; and improved provision of data to support policy development.<sup>87</sup>

---

84 <http://www.whitehouse.gov/sites/default/files/omb/inforeg/for-agencies/informing-consumers-through-smart-disclosure.pdf>

85 <http://www.smh.com.au/business/place-for-publicprivate-partnerships-in-cyberspace-20130328-2gxdp.html#ixzz30Ahdm1ui>

86 Australian National Data Service (ANDS). Available <http://www.ands.org.au/about-ands.html>

87 <http://www.ands.org.au>

#### 4.4 Encouraging open private data

In certain circumstances, and given certain information architectures, there can be strong reasons for private firms to donate digital assets that they have built to the digital commons. The canonical example is open source software. Here firms or individuals using the software freely have an incentive to fix bugs or add features to the software to enhance its functionality for them. Then, once they have written the code, they have an incentive to donate it back to the project. That is because the software is continually updated with any number of improvements to it, so the community that uses the software has an interest in using the updates. But if someone has built a bug fix or new feature into the software, they do not want to have to code it back into the updated software when the next distribution is released. So it is generally in the interests of all the users of the software to donate their improvements back to everyone else. Everyone mutually gains from the process.

Something similar can exist in the case of private data where that data acquires greater usefulness when combined with other data. This will often be the case on some platform. It is on this basis that the Sense-t project is being funded by the Federal and Tasmanian governments, which integrates data freely provided by private farmers and others from existing sensors around Tasmania into a 'platform' on which the data can be endlessly used, reused and re-purposed, and on which 'apps' can be built to add value to the data. This is already being used to optimise agricultural performance in a range of industries such as viticulture, aquaculture and water management, but it can be used in virtually all agricultural industries and industries requiring land management and spatial and environmental information.

#### Building data sharing around consumer consents

The UK government's "Midata" strategy<sup>88</sup> is built on the right citizens of the UK have to access data that UK corporations may hold on them (a right that is not provided by many other countries including the US). This may be assembling a powerful means by which value added services can be built and then marketed to those who share their data, with users remaining in complete control as to what data they release to which parties for what purposes. It is easy to imagine brokers emerging managing customers' private data so as to maximise the flow of desired digital services to citizens. The UK legal architecture and government activity around Midata seems likely to open up powerful new opportunity for building value from data which is built from the ground up on citizens own rights to, and control of their data.

As McKinsey Institute puts it, the use of open data:

*Requires governments to play a central role by developing and implementing policies to mitigate consumer and business concerns about the misuse of open data and to help set standards that will allow the potential economic and social benefits to materialize.*<sup>89</sup>

Nevertheless, however promising such an approach seems, as recent events in the UK demonstrate, it is imperative for those seeking to engineer a world which is more alive to the value of open data and data sharing to retain citizen trust by ensuring that citizens feel properly informed and confident that their privacy will be protected as a foundation of the system.

---

88 <https://www.gov.uk/government/news/the-midata-vision-of-consumer-empowerment>

89 [http://www.mckinsey.com/insights/business\\_technology/open\\_data\\_unlocking\\_innovation\\_and\\_performance\\_with\\_liquid\\_information](http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information)

#### **4.5 The open data opportunity and the G20: What governments need to do – individually and collectively**

The open data opportunity spans government data (PSI), data arising from publicly funded research, and private, both business and personal, data. Hence the opportunity spans:

- Direct open data policies focusing on making more government data openly available;
- Supporting and prosecuting policies focusing on opening publicly funded research data (e.g. through research funding agencies); and
- Establishing a regulatory and policy environment conducive to enabling the maximisation of opportunities to open private sector data.

Policies in these areas have both national and international dimensions. Thus, there is an important role for international forums, such as the G20, including sub-groups, such as B20. A series of case studies presented in the Annexes show both the potential impacts of open data, and reveal much about the issues, barriers and policy-related opportunities that arise both at the national and international levels. The following sections summarise those case studies with a particular focus on their policy lessons for governments acting individually and collectively.

##### **G20 Agenda Item – Annex 1 – Trade**

There are many difficulties associated with obtaining and using government trade-related data from multiple jurisdictions and regions. There can be price barriers, as there is a mix of free and paid data, and the data can require related processing/consulting services from the data providers. Sometimes even small access fees involve substantial transaction costs. How much easier if all government data of this kind were freely available, not just nationally, but across all G20 jurisdictions?

Licensing conditions can also present problems, and it can take many hours ensuring understanding of and compliance with licenses from different countries, different legal jurisdictions, in different languages, etc. How much easier if all G20 governments reported to similar standards and used standard and permissive licensing, such as creative commons?

Data can be made available and/or be delivered in different file formats, using different media, as well as involving a variety of national commodity classifications, and exhibiting varied quality. How much easier if standard formats and data schema were used throughout the G20 and beyond?

The data that each country generates is a potential global public good as overwhelmingly each party to new trading patterns gains and the data helps the global economy search out its most efficient configurations of production, exchange and consumption. How much easier if open data policies were harmonised across the G20 countries and beyond?

##### **G20 Agenda Item – Annex 3 – Fiscal and monetary policy**

Exploring the case of using open data to better manage economic cycles, we suggest that it could contribute AUD 3.6 billion per annum to Australia's economy alone.

Governments are in possession of rich real time data that remarkably, even today, is not routinely accessible in a timely fashion to macro-economic managers or the wider world. Government agencies, such as the Treasury and the Reserve Bank of Australia, use sophisticated models to forecast and to assess the impact of their policy choices. These models are not well documented publicly. However, in some countries authorities are beginning to release the models they are using. The US Federal Reserve has released one of its models of the US economy which it uses internally. Similarly, the Australian Treasury is going through a process of releasing the econometric equations/approaches that lie behind its forecasts. There would be

enormous benefit from all G20 countries following this lead, allowing, for example, researchers to compare approaches and identify possible amendments and improvements.

There are already private endeavours to develop and make better use of real time data, but Governments could move much faster in taking advantage of this trend. More and more financial services are being delivered via software delivered from 'the cloud' as a service. In such cases service providers can aggregate the data to provide real-time snapshots of the state of the economy. Thus, where the vendors of 'software as a service' make software services available 'in the cloud', governments could seek out arrangements with them to supply anonymised data to government agencies for various reasons including to 'take the pulse' of the economy in real time.

Gruen has also suggested that governments could hold a competitive tender between service providers to provide on-line services such as accounting and other software to businesses free of charge.<sup>90</sup> This would bring the efficiency gains of pricing business software down to its marginal cost whilst maximising the extent to which data between different users of such services were compatible. This could offer the means by which governments could ensure a steady stream of highly representative, anonymised, real-time data on the performance of the economy.

#### **G20 Agenda Item – Annex 4 – Anti-Corruption**

The cost of corruption in developed countries is significant. Johnson et al. (2010)<sup>91</sup> examined differences in the growth and corruption among US states finding evidence consistent with corruption having a negative and significant effect on growth and leading to lower levels of capital investment. Open data may directly contribute to reduced corruption by increasing the likelihood corruption will be detected. For example, Eaves (2010) describes how open data provided by Canada Revenue Agency (CRA) helped to 'expose one of the biggest tax frauds in Canada's history'.<sup>92</sup>

E-government is recognised as having the potential to combat corruption by improving the enforcement of rules, lessening the discretion of public officials and increasing transparency. Krishnan et al (2012) found e-government maturity negatively contributes with corruption.<sup>93</sup>

Thus in addition to its other economic and social benefits, the proactive release of machine readable data in interoperable formats enabling it to be combined and compared with other data can provide a powerful means of fighting corruption. To this end Governments should work on not just the release of data, but also collaborating on harmonised technical standards that permit the tracing of international money flows – including the tracing of beneficial owners of various commercial entities, and the comparison and reconciliation of transactions across borders.

#### **G20 Agenda Item – Annex 5 – Employment**

Open data has potential to transform labour markets through transmission of information on work-place

- 
- 90 Gruen, N, 2014, "Government as Impresario: Emergent public goods and public private partnerships 2.0", 2014, Presentation to the Berkman Centre for Internet and Society, Cambridge, Massachusetts, January 14. <http://cyber.law.harvard.edu/events/luncheon/2014/01/gruen>
- 91 Johnson, N. D., LaFountain, C. L., & Yamarik, S. (2011). Corruption is bad for growth (even in the United States). *Public Choice*, 147(3-4), 377-393.
- 92 Eaves, David (2010). Case Study: How Open data saved Canada \$3.2 Billion, Available <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>.
- 93 Krishnan, S., Teo, T.S., & Lim, V.K. (2012) E-Government Maturity, Corruption, And Eco-Efficiency. *E-Government*, 7, 15-2012.

conditions to prospective employees. Consider the impact of higher quality information on employment conditions being available to job-seekers. This would improve the matching of employees to organisations improving job-satisfaction, and so, productivity. A longer-term benefit would be the impact on workplace quality as firms, more concerned than ever with their reputation as employers, seek to improve workplace conditions. This in turn should also encourage greater work-force participation.

Governments cannot wave a wand and banish such asymmetric information problems away. But they already have some tools with which they can begin to make progress. They can release their own data on employee opinions on the quality of various government agencies. And they can help play a role in the emergence of standards – for instance by convening employers to agree on such standards, releasing their own data against them and deploying their statistical agencies to survey a random sample of firms to provide information according to which others can calibrate their own data – to find where a workplace is amongst its peers. Once such standards exist, it becomes in the interests of the best performing firms to collect and publish data according to that standard (See Box 9).

### **G20 Agenda Item – Annex 6 – Energy**

Annex 6 explores the potential role of open data in the energy sector, particularly in the areas of generational and distributional optimisation, infrastructure design and development and efficiency of use.

An example of the latter is the US energy company Opower,<sup>94</sup> which uses open data from consumers and suppliers to enable cost savings through optimising energy plans, peer level comparison and strategies to change consumption behaviour.

Enabling change in the energy sector will require data-empowered consumers and consumer advocate organizations with the expertise to offer more energy-efficient strategies and solutions. While smart meter technologies will provide much needed consumption data, it will be vital that consumers have free and open access to all of their consumption data, something that policy can play a role in. Equally, ensuring industry competition requires open data to provide contestability. Provision of an adequate open data framework represents a crucial part of making sure that industry transformation is based on data-driven evidence. Energy efficiency is a global, not just national issue, and there is an opportunity to standardise energy use data and share it internationally, to enable the identification of best practices from as wide a field as possible. Comprising the largest national economies, the G20 could play a leading role, not just catalysing the opening of data, but also in harmonising and standardising regulations and requirements, to change the isolated pools of national data into worldwide oceans of information that could provide the foundation for substantial energy efficiency improvements.

### **G20 Agenda Item – Annex 7 – Infrastructure**

The very high cost of developing new infrastructure to meet growing demands leads to two key infrastructure challenges:

- maximising the value obtained from existing infrastructure, and
- ensuring the efficient development of new infrastructure.

Open data can assist with both of these.

Improved information made available as part of an open data focus can help improve utilisation in a number of ways (e.g. congestion alerts, travel time estimates, etc.). Improved information may also help other parties, such as those providing services associated with the infrastructure. For example, knowledge on road con-

---

94 [www.opower.com](http://www.opower.com)

gestion may be used by a port operator for planning purposes. A further benefit of improved utilisation information is that it may provide a foundation to facilitate other regulatory reform (for instance congestion pricing).

Greater transparency and greater capture and release of data may play a key role in infrastructure spending, through improved:

- scrutiny of key infrastructure investment decisions;
- analysis of costs and benefits;
- implementation planning and coordination, and
- data can also aid procurement process.

Current public sector infrastructure asset information is fragmented across industry and government sectors. In many countries, local, state and federal governments maintain disparate asset data. The fragmentation leads to inefficiencies. Exposing current "business-as-usual" asset data would represent a significant first step in understanding gaps. It would begin to provide new insights. Doing so in a consistent and harmonised way across G20 countries could provide a wealth of information enabling greater efficiency in infrastructure development and use.

## 5 Conclusion

Our work suggests that an even stronger embrace of the open data agenda can generate substantial economic gains for Australia and for its partners in the G20. Indeed this report demonstrates that those gains are sufficiently large that they could make a substantial contribution to the G20's 2% growth target. Indeed, one could make some more optimistic assumptions that were still not unreasonable and argue that they might achieve the growth target on their own.

Policies facilitating open data have both a national and an important international scope. Indeed, given the increasing intensity of globalisation, fully realising the open data opportunity is going to require cooperative and coordinated international action in such areas as standardisation and harmonisation of national regimes. Consequently, there is an important role for international forums, such as the G20.

But there is another perhaps even more compelling argument. It essentially involves making better use of an existing resource. Virtually any other micro-economic reform generating comparable gains, would involve hard choices, visible job losses and commensurate political difficulties for those introducing the change. By contrast, the kinds of changes envisaged in this report – involving getting more from our existing data, improving our management of the cycle and the matching of people to workplaces – are all things that would be welcomed by the community because of the obvious good that they would do.



## Annexes Case studies of the impacts of open data

The following case studies demonstrate how open data can create value. These examples have been selected for their alignment with the themes of the G20 Agenda (Figures 3 and 12). The case studies and examples are also used to demonstrate the many innovative ways in which governments can act to create value through open data, both individually and collectively through such forums as the G20.

### **Annex 1 G20 Agenda Item – Trade**

TradeData International (<http://www.Tradedata.net>) offers one example of how open data can improve the efficiency of international trade and facilitate economic development in developed and emerging regions.

#### **History and operations**

TradeData International has been operating since 1995, initially as a part of Victoria University (Melbourne), before spinning off as a separate business in 2001. The business focuses on collecting and analysing trade data for around the world (currently from 100 countries), and providing commercial information and research services to governments and industry. The data are based on customs returns, and updated monthly, leading to an extensive and rich data source.

The data are sourced predominantly from national statistical agencies, so this is government data. TradeData then adds value to the government data by working to harmonise trade categories and correct errors in the data, before adding it into its database, upon which it then builds its analysis and consulting services business.

The majority (80-90%) of TradeData's clients are private sector customers, mostly in manufacturing. These are both subscriber and ad hoc purchasers. The remaining 10-20% of clients are public sector and development/aid organizations.

#### **Benefits**

TradeData's data and analytical expertise provides a range of benefits including: helping identify and so facilitate trade and development opportunities, reducing costs and risks of trading for its clients, and improving the efficiency of international trade.

For firms, TradeData can analyse competitor activities in current or potential markets, search the world for potential export opportunities, assess historical firm performance relative to the performance of the total sector, and quantify market entry conditions and thus assist the firm in developing a market entry or export strategy. For governments, TradeData can assist in the formulation of national and regional trade and development strategies, and monitor performance and progress. Governments often target selected overseas markets for export promotion, sometimes sending a representative to promote local industry export capabilities. TradeData can assist by providing detailed market analysis of most countries around the world on a very specific, product-by-product basis.

TradeData regularly undertakes specific product or sector analyses for major international development agencies like the Asian Development Bank, International Finance Corporation, World Bank, AusAID and others. Typically, these studies help to understand the dynamics of the world market and advice how a country or region may change their product mix and/or strategy to increase the average value and volume of their exports.

Leading economists at Harvard led by Ricardo Hausmann have pioneered a new approach to economic development showing how frequently new development and export opportunities cluster around areas that

closely resemble existing areas of know-how.<sup>95</sup> This explication of the “adjacent possible” provides a powerful tool with which to conceptualise promising development opportunities for an economy. The TradeData International database offers a complementary tool for identifying development and trade opportunities in an economy.

### Examples

TradeData prepared an analysis of trade opportunities for the State Government of Victoria, in which it combined analysis of the state’s industry strengths with detailed analysis of the prices at which Victoria’s exports could be landed in various ports around the world compared to the actual landing prices of those commodities, identifying not only with what products, but also where Victoria’s producers would be most competitive. This enabled a targeted export promotion campaign based on informed analysis, rather than guesswork and wishful thinking.

In another case, TradeData became involved in a development program in Pakistan focused on low-income farmers.<sup>96</sup> The viability of smallholders is crucial for the sustainability of Pakistan’s agriculture, so the aim is to develop options, and evaluate and define enabling policies to improve the livelihoods of smallholders in the dairy, citrus and mango sub-sectors of the Punjab and Sindh provinces of Pakistan. As a part of the project, TradeData has undertaken analyses of global market opportunities for smallholder products (e.g. mango) and determining export strategies based on seasonal demand for fresh mango and mango-related products.

### The open data opportunity

Over the years, TradeData has encountered most, if not all, of the difficulties associated with obtaining and using government data from multiple jurisdictions and regions.

There can be price barriers, as there is a mix of free and paid data, and the data can require related processing/consulting services from the data providers. TradeData reports paying over time many hundreds of thousands of dollars in data fees. Sometimes even small access fees can involve substantial transaction costs. How much easier if all government data of this kind were freely available?

Licensing conditions can also present problems, and TradeData staff have spent many hours ensuring understanding and compliance with licenses from different countries, different legal jurisdictions, in different languages, etc. How much easier if all governments reported to similar standards and used standard and permissive licensing, such as creative commons?

The data that each country generates is a potential global public good as overwhelmingly each party to new trading patterns gains and the data can help the global economy search out its most efficient configurations of production, exchange and consumption. How much easier if open data policies were harmonised across the G20 countries and beyond?

Data can be made available and/or be delivered in different file formats, using different media, as well as involving a variety of national commodity classifications, and exhibiting varied quality. How much easier if standard formats and data schema are used?

In many ways, TradeData was ahead of its time, and things have improved enormously on all these fronts since 1995. But there are still costs and barriers that could be lowered, possibly enabling TradeData to reduce prices and expand its business and facilitate the emergence of business like it.

---

95 <http://www.hks.harvard.edu/centers/cid/publications/featured-books/atlas>.

96 See <http://www.cfses.com/cses/news.htm>

It is impossible to estimate exactly how much that might be worth, but it is easy to imagine how it might work. The Victoria and Pakistan examples are indicative. With global exports worth some USD 25 trillion per annum, of which merchandise exports are around USD 20 trillion,<sup>97</sup> a very small efficiency improvement in trade could be worth a lot.

For example, an expansion of world merchandise trade by just one-tenth of 1% would be worth USD 40 billion per annum at today's prices. But, in the longer term, bringing conflict ridden and emerging regions into global markets, as in the Pakistan example, may be worth even more.

The potential efficiency improvement is very large. Research on the costs of trade has highlighted that information barriers impose a substantial cost. The empirical research on the costs of trade suggests that for a representative rich country, information barriers are equivalent in impediment to a 6 per cent ad-valorem tax<sup>98</sup> (that does not provide tax revenue). Given this, a potential benefit of open-trade data of one tenth of information barriers seems plausible. This amounts to 0.6 per cent of merchandise trade, which for Australia would be equivalent to around AUD 1.6 billion or 0.10 per cent of GDP.

## **Annex 2 G20 Agenda Item – Finance**

In 2001, Sirca identified that an archive of global financial market data comprising tick-by-tick data for all world markets held by Thomson Reuters had value to finance and economics scholars around the world who were seeking to better understand market micro-structure on a global scale. At this time, the archive had no perceived value to business or government.

Established as a cooperative between Australian and New Zealand universities and initially funded by public research infrastructure grants with co-investments from the university and business sectors, Sirca was able to build the world's first online global financial market data service for academic research.

In 2005, having proved the value of such a "big data" service to academia, Sirca was able to further enhance and adapt this service for business and government use and attract major investment from Thomson Reuters (TR). For fund managers, investment banks and hedge funds, this unique data archive enables cross-market algorithmic trading and compliance activities. For government regulators and central banks, the data provides a means to gain insights into trading behaviour for market surveillance and improved policy design and implementation.

Today Sirca's facility ingests over 2 million records per second from all major financial markets and is the world leader in re-purposing this high frequency data for research, business and regulatory use. The Tick History data service now supports analytics at almost 600 universities, banks, funds managers and regulators world-wide through an innovative "analytics-ready" online platform.

Sirca's Tick History service has had significant outcomes in both public and private spheres. The service has created a whole new paradigm of cross-market micro-structure research enabling over 1,500 academic research publications, which in turn have been cited by around 10,000 papers since inception. In addition, the service has become the leading data source for the development and testing of high frequency trading, compliance, and regulatory analytics globally. The Sirca Group activities have also created over USD 1 billion in industry revenues and over 600 direct jobs from its own growth as well as spin-off enterprises.

## **Annex 3 G20 Agenda Item – Fiscal and Monetary Policy**

In 1987, Lucas canonically proposed that the cost of the economic cycle imposed negligible costs on con-

---

97 See [http://www.wto.org/english/news\\_e/pres13\\_e/pr688\\_e.htm](http://www.wto.org/english/news_e/pres13_e/pr688_e.htm)

98 Anderson, J. E., & van Wincoop, E. (2004). Trade Costs. *Journal of Economic Literature*, 42(3), 691-751..

sumers who bear only the costs of small variations in consumption. However Lucas's intervention occurred near the zenith of confidence that recessions were essentially temporary affairs, that economies bounced back from recessions with little long-term economic cost as they returned to the growth path from which they had temporarily deviated (Ball et al., 2014).<sup>99</sup>

In today's world in which the effects of the 2008 global financial crisis continue to traumatise the world economy, this idea of the relative benignity of the economic cycle seems far fetched. As Ball, DeLong and Summers put it:

*The textbook model of short-term recessions is contradicted by research based on broader international data. International Monetary Fund (IMF) studies, such as the 2009 World Economic Outlook, that look at post-World War II financial crises have found that essentially all of the output decline associated with a typical crisis persists for at least seven years, and little or none of the shortfall relative to the pre-crisis trend is recovered within that time span.*<sup>100</sup>

Among the things that Lucas's model assumes away are a range of phenomena that were always at the heart of people's concern with the costs of recessions such as the heterogeneous impact on different workers, the impact on long-term unemployment and the extent to which recessions blight the lives not just of those rendered unemployed, but also of those entering the workforce during recession.<sup>101</sup> Taking some of these things into account, Galí et al. (2002) estimate the costs of the economic cycle at over an order of magnitude higher than Lucas at 1.35% of consumption.<sup>102</sup> This looks like an underestimate given that it is calibrated from data that includes 'the Great Moderation' but not including the Great Recession still upon us.

Though, to be conservative, we will not take them into account in this exercise, it can be noted parenthetically that there are numerous other potentially large costs of the cycle. Thus the cost of asset bubbles on consumption smoothing has substantial costs (Chauvin et al., 2011).<sup>103</sup> Likewise, as Grant and Quiggin observe, contra Lucas, "the larger the risk premium on equity, the larger the marginal value of recession-state income and the greater the welfare cost of recessions (2006, p 262).<sup>104</sup> Further, and remarkably, recession-born start-ups tend to remain persistently smaller on average than start-ups born in more prosperous times, even when the aggregate economy recovers leaving an "increasingly large footprint on the macro-economy as start-ups age" (Sedláček, 2014).<sup>105</sup>

---

99 Ball, Laurence, DeLong, Brad, and Summers, Larry, 2014. "Fiscal Policy and Full Employment", *Mimeo*, April 2, [http://www.pathtofullemployment.org/wp-content/uploads/2014/04/delong\\_summers\\_ball.pdf](http://www.pathtofullemployment.org/wp-content/uploads/2014/04/delong_summers_ball.pdf)

100 Ball, Laurence, DeLong, Brad, and Summers, Larry, 2014. "Fiscal Policy and Full Employment", *Mimeo*, April 2, [http://www.pathtofullemployment.org/wp-content/uploads/2014/04/delong\\_summers\\_ball.pdf](http://www.pathtofullemployment.org/wp-content/uploads/2014/04/delong_summers_ball.pdf)

101 Yellen and Akerlof, 2006, pp. 4-6 for an enumeration of four of Lucas's assumptions – all tending to reduce the measured cost of the economic cycle.

102 Galí, Jordi, Mark Gertler, and J. David López-Salido, 2002. "Markups, Gaps, and the Welfare Costs of Business Fluctuations," NBER Working Paper 8850, Cambridge, MA, March. See also Yellen and Akerlof, 2006, pp. 4-6 for an enumeration of four of Lucas's assumptions – all tending to reduce the measured cost of the economic cycle. Krusell, P., Mukoyama, T., Şahin, A., & Smith Jr, A. A. (2009). Revisiting the welfare effects of eliminating business cycles. *Review of Economic Dynamics*, 12(3), 393-404.

103 Chauvin, Kyle, David I. Laibson, and Johanna Britta Mollerstrom. 2011. Asset bubbles and the cost of economic fluctuations. *Journal of Money, Credit and Banking*, 43(S1): 233–260.

104 Grant, Simon, and Quiggin, John, 2006. "The Risk Premium for Equity: Implications for Resource Allocation, Welfare and Policy", *Australian Economic Papers*, Vol. 45, No. 3, pp. 253-268, September.

105 Sedláček, Petr and Sterk, Vincent, 2014, "The Growth Potential of Startups over the Business Cycle", March 29, 2014, [http://www.psedlacek.com/Documents/Working/startups\\_20140328.pdf](http://www.psedlacek.com/Documents/Working/startups_20140328.pdf)

### Using open data to better manage the economic cycle

There are three ways in which a more open approach to data can assist in the management of the economic cycle. The first two, explored in the next section, make better use of existing data while the third involves the exploitation of the new possibilities of the new information technology infrastructure of the internet age to build additional 'big data' resources to understand and manage the economic cycle.<sup>106</sup>

### Making better use of existing data

Government is in possession of rich real time data that remarkably, even today, is not routinely accessible in timely fashion to macro-economic managers or the wider world. This includes most particularly real time GST, income and company tax payments. Secondly, government agencies such as the Treasury and the Reserve Bank use sophisticated models to forecast and to assess the impact of their management of policy settings. These models are not well documented publicly. However, in other countries, authorities are beginning to release the models they are using. Thus, for instance, on the 3rd April 2014, the US Federal Reserve announced the release of one of its models of the US economy it used internally.<sup>107</sup> As John Taylor observed praising the Fed on this move "The ability to replicate is essential to good research, and the same is true of good policy research".<sup>108</sup> The Australian Treasury is going through a process of releasing the econometric equations/approaches that lie behind their forecasts. This openness could and should be embraced in a more programmatic way by Australian and other G20 governments.

### 'Big data' resources to better understand and manage the economic cycle

Another era is approaching in which much of the data used to manage the macro-economy will be gathered in real time and used either in real-time or following a processing lag of a few days or even hours rather than the months that are customary today. Even the use of currently available real-time financial data appears to improve economic forecasting (Andreou, et al., 2013). But the future holds far more exciting prospects than the simple use of financial data from share markets and similar institutions.<sup>109</sup>

---

106 It should be noted here that not only are we failing to make optimal use of the data we already have, monthly Reserve Bank (RBA) Board meetings are timed with almost exquisite disregard for the timing of major data releases. "Four of [the RBA's] 11 interest rate decisions each year are made three months after the previous quarterly national accounts release and the day before the next one. Further, abandoning the practice of assuming that the economy can wait for board members to enjoy their summer break would generate some gains and could be agreed upon immediately." See, Gruen, N. 2012. "RBA board needs to chart new course", "The Sydney Morning Herald and the Age", June 14, <http://www.smh.com.au/business/rba-board-needs-to-chart-new-course-20120613-20ahp.html>

107 <http://www.federalreserve.gov/econresdata/frbus/us-models-about.htm>

108 <http://economicsone.com/2014/04/05/transparency-for-policy-works/>

109 Andreou, E., Ghysels, E., & Kourtellis, A. 2013. Should macroeconomic forecasters use daily financial data and how?. *Journal of Business & Economic Statistics*, 31(2), 240-251. [http://www.ucy.ac.cy/hermes/documents/ConferencesEvents/Macro\\_MIDAS\\_Nov\\_2\\_2010\\_EG.pdf](http://www.ucy.ac.cy/hermes/documents/ConferencesEvents/Macro_MIDAS_Nov_2_2010_EG.pdf)

### Box 12 Uncovering public and private value from private data

[PeopleLikeU.com.au/](http://PeopleLikeU.com.au/) is a public demonstrator of private consumer spending habits in Australia based on transaction data from UBank, an NAB Subsidiary. It enables users to better understand spending, saving and specific purchase preferences based on demographic data filters. The site has been built to be of use to citizens who may be curious about their own spending habits compared with peers. Thus if users provide their age, place of residence and income range, they can browse to see what people like them are spending their money on – where they go to holiday, what restaurants they prefer and so on.

This kind of data should be of considerable use to firms doing market research and seeking market intelligence about the sizes of various market segments, and real-time trends. It should also improve economic forecasting and so management of the economy generally and regional development.

Source: "Open Government Lighthouse Initiatives", provided by Sirca Pty Ltd

There are already private endeavours to develop and make better use of real time data, but governments have it within their power, and should accelerate this trend. More and more financial services are being delivered via software delivered from 'the cloud' as a service. In such cases service providers can aggregate the data to provide real-time snapshots of the state of the economy. Thus, where the vendors of 'software as a service' make software services available 'in the cloud', governments could seek arrangements with those software providers to supply anonymised data to them for various reasons including for the purpose of 'taking the pulse' of the economy in real time.

Gruen (2013) has also suggested that governments could hold a competitive tender between service providers to provide on-line services such as accounting and other software to businesses free of charge. This would bring the efficiency gains of pricing business software down to its marginal cost whilst maximising the extent to which data between different users of such services were compatible. This could offer the means by which governments could ensure a steady stream of anonymised real-time data on the performance of the economy.<sup>110</sup>

### Box 13 Premise Data and the real-time measurement of inflation

The world of economic data may be headed for a substantial revision.

Premise Data is the first to create real-time inflation data using hundreds of people snapping photos of store shelves. Software automatically tags the location of the products down to the individual store. In contrast, the Labor Department collects price data by dispatching its staff to collect product prices once a month.

Technological innovations such as efficient supply networks and just-in-time delivery "have completely altered human economic activity," said David Soloff, founder and chief executive of Premise. "Yet the indicators people are looking to are caught in a different era."

110 Gruen, N, 2014, "Government as Impresario: Emergent public goods and public private partnerships 2.0", 2014, Presentation to the Berkman Centre for Internet and Society, Cambridge, Massachusetts, January 14. <http://cyber.law.harvard.edu/events/luncheon/2014/01/gruen>

Even when the government is open and operating, reports on inflation, employment and consumer spending—generally based on monthly snapshots—are released with a lag of weeks or months. “There’s still a Cold War element to our statistics,” said Alan Krueger, a Princeton University professor and former top White House economist, who [is] a Premise adviser.

Hal Varian, Google Inc.’s chief economist and a Premise adviser, said the data could give government officials insight into developments that can stir up their populations. Prices for popular food items—bread in the Middle East, corn in Mexico, or pork in China—could be tracked well ahead of popular unrest.

“All these things are sensitive from a political point of view,” Mr. Varian said. “Having up-to-date information is quite valuable.”

Source: *Wall St Journal*, Oct 15, 2013.<sup>111</sup>

### **Estimating the benefits of enhanced fiscal and monetary policy**

We take the costs of the business cycle as moderated using existing macro-economic policy to be 1.35% of consumption or a little under 1% of GDP. As indicated, this is a very conservative figure given that this number was calibrated from economic fluctuations available during the Great Moderation and before the Great Recession struck with the Global Financial Crisis of 2008. Now the benefits of such a reduction in severity are non-linear. The Taylor rules used by Central Banks have a quadratic loss function implying that the economic cost of recessions is proportional to the square of their severity (which will be a function of their amplitude and their duration). This is good news as it means that relatively modest improvements in economic management produce disproportionate gains.

We think it reasonable to suggest that the measures we have proposed would reduce the severity of fluctuations by around one-eighth and that this would reduce the economic cost of those recessions by nearly double that! We propose that better use of existing resources – opening up existing tax and other data and opening up official forecasting models and methods could further improve macro-economic management and reduce the severity of future downturns by 5%. Nonlinearity would then deliver a welfare gain or avoided economic costs of nearly 10% (strictly 9.75%) of the cost of existing fluctuations.

We think that the third advance outlined above – the harnessing of ‘big data’ to the management of the cycle – should improve macro-economic management somewhat more than the better use of existing resources. The development and use of real-time big data should enable more precise and confident appraisal of the timing, magnitude and targeting of measures based on existing economic theory. It should also help us refine that theory somewhat. And so we propose that this would reduce the remaining severity of recessions by a further 7.5%. The benefit of this extra 7.5% improvement shaves an additional 13.7% of the cost of economic fluctuations,<sup>112</sup> for a total reduction in the economic cost of the cycle of 23.4%.

111 Real-Time Economic Data Could Be a Game Changer, *Wall St Journal*, Oct. 15, 2013 <http://online.wsj.com/news/articles/SB10001424052702304330904579135832100578954>

112 Note that, relatively speaking, the magnification of the gain delivered by non-linearity is not quite as much as that delivered by the first five percentage point reduction in the severity of fluctuations as non-linearity means that the greater the gains already achieved, the fewer can be captured by further improvements. As discussed above, if we could reduce fluctuations by two-thirds, we would already have removed eight-ninths of their cost.

**Table 5 Economic benefits from better use of open data in managing economic fluctuations (AUD billions per annum)**

<i>Total GDP</i>	1,554
Consumption	1,136
Existing cost of fluctuations (1.35% of consumption)	15.3
9.75% of existing cost of recessions	1.50
13.7% of existing cost of recessions	2.10
Total benefits of open data	3.60

Source: *Lateral Economics' analysis.*

Given that the existing cost of fluctuations is assumed to be 1.35% of total final consumption (which is around AUD 1.5 trillion dollars), the total value of the benefits of open data in improving fiscal and monetary policy can be conservatively estimated at around AUD 3.6 dollars per annum or 0.23% of GDP per annum.

#### **Annex 4 G20 Agenda Item – Anti-corruption**

Corruption can be broadly defined as acts in which the power of public office is used for personal gain in a manner that contravenes the rules of the game (Jain 2001).<sup>113</sup> There is a rich economic literature on corruption and the costs of corruption (see for example, Jain 2001 and more recently Dreher and Herzfeld 2005).<sup>114</sup>

Recent work has emphasised the cost of corruption in developed countries to be significant. Johnson et al. (2010)<sup>115</sup> examined differences in growth and corruption among US states. They found evidence consistent with corruption (as measured by corruption convictions) having a negative and significant effect on growth in U.S. states and leading to lower levels of capital investment. Their analysis highlighted the significance of corruption, indicating at the margin that a one standard deviation change in the level of corruption leads to a fall of 0.19 to 0.5 percentage points in annual growth rates.

More recently a European Commission (2014) report on anti-corruption in the EU<sup>116</sup> estimates of the costs of corruption in the EU to be around 1% of GDP, equivalent to around EUR 120 Billion.<sup>117</sup> The report also cites survey evidence that suggests that most people in European countries think corruption is widespread in their country and over a quarter think they are personally affected by corruption.

#### **Benefits of open data in reducing corruption**

Open data, as part of open government, can help to reduce the costs of corruption. Open data can also be used to promote transparency and drive accountability, thereby reducing the economic loss due to corruption. Open data may contribute to a reduction in corruption and benefits to economies in a number of ways.

113 Jain, A. K. (2001). Corruption: a review. *Journal of economic surveys*, 15(1), 71-121. At [http://darp.lse.ac.uk/PapersDB/Jain\\_\(JES01\).pdf](http://darp.lse.ac.uk/PapersDB/Jain_(JES01).pdf), accessed 27 Apr, 2014.

114 Dreher, A., & Herzfeld, T. (2005). The economic costs of corruption: A survey and new evidence. *Public Economics*, 506001.

115 Johnson, N. D., LaFountain, C. L., & Yamarik, S. (2011). Corruption is bad for growth (even in the United States). *Public Choice*, 147(3-4), 377-393.

116 European Commission (2014). Report from the Commission to the Council and the European Parliament: EU anti-corruption report. COM(2014) 38 final Brussels, 3.2.2014.

117 The costs of corruption are not easily measured. The cited figure is based on estimates made by other institutions which suggest that costs of corruption amounts to 5% of GDP worldwide.



It may directly contribute to reduced corruption by increasing the likelihood corruption will be detected. For example open and interoperable data could allow citizens to follow — and consequently shape — the flow of public money across borders thereby reducing the risk of misuses of public funds and resources. For example, the opening of data is enabling journalists to expose corruption and financial mismanagement on the revenues from extractive industries.<sup>118</sup> Open data would facilitate governments working together to tackle global issues such as tax evasion, the use of anonymous shell companies for money-laundering and other corrupt practices.

There are indirect benefits as well. Increased rates of detection would reduce the returns to corruption and thereby deter corruption occurring in the first place. Reducing the risk of corruption could enhance competition and improve business investment. For example, it would enable governments to work together to create a level playing field for business, thereby reducing the returns to corruption. By improving the information businesses have to assess risks and opportunities in different markets so that they can make better investment decisions.

Due to the measurement difficulties, there is limited empirical research that relates open data and corruption. However, there is sufficient evidence to suggest that the contribution of open data to reducing corruption can be significant. A simple demonstration of how open data can work against corruption is provided in a case study by Eaves (2010)<sup>119</sup> (see Box 14). Eaves describes how open data provided by Canada Revenue Agency (CRA) helped to 'expose one of the biggest tax frauds in Canada's history'. Eaves makes the analogy of detection of corruption with the detection of computer software problems, which is said to follow what is known as Linus' Law; a law which states: "given enough eyeballs, all bugs (problems) are shallow."

---

118 See for example, 'Covering the Extractives Industry: Big Data, New Tools, and Journalism' By: Anya Schiffrin and Erika Rodrigues, March 5, 2014. At <http://gjin.org/2014/03/05/covering-the-extractives-industry-big-data-new-tools-and-journalism/>

119 Eaves, David (2010). Case Study: How Open data saved Canada \$3.2 Billion, At <http://eaves.ca/2010/04/14/case-study-open-data-and-the-public-purse/>, accessed 27 Apr, 2014.

#### **Box 14      How open data uncovered a tax fraud**

Eaves (2010), presents an enlightening case of how open data contributed to the exposure of tax fraud in Canada:

*[...] a colleague was asked by a client to do an analysis of the charitable sector in Toronto. Considering it a simply consulting project, my colleague called the Canada Revenue Agency (CRA) and asked for all the 2005 T3010s – the Annual Information Returns where charities disclose to the CRA their charitable receipts and other information – in Toronto.*

The research uncovered illegally operating charities alone sheltering very large sums:

*Indeed, newspapers later confirmed that in 2007, fraudulent donations were closer to a billion dollars a year, with some 3.2 billion dollars illegally sheltered, a sum that accounts for 12% of all charitable giving in Canada.*

Eaves (2010), concluded:

*This means that, at best, government data – information that should be made more accessible and open in an unfettered and machine readable format – helped reveal one of the largest tax evasion scandals in the country's history. But if the CRA was already investigating, scrutiny of this data by the public served a different purpose – helping to bring these issues out into the open, forcing CRA to take public action (suspending these organizations' right to solicit more donations), sooner rather than later. Essentially from before 2005-2007, dozens of charities were operating illegally. Had the data about their charitable receipts been available for the public's routine review, someone in the public might have taken notice and raised a fuss earlier. Perhaps even a website tracking donations might have been launched. This would have exposed those charities that had abnormally large donations with few programs to explain them. Moreover, it might have given some of the 100,000 Canadians now being audited a tool for evaluating the charities they were giving money to.*

*Source: Abridged from Eaves (2010) 'Case Study: How Open data saved Canada \$3.2 Billion'.*

#### **What governments need to do**

The benefits of open data in addressing corruption requires proactive action in both releasing data and making it machine readable according to interoperable standards enabling it to be combined and compared with other data. To this end Governments should work on not just the release of data, but also collaborating on harmonised technical standards that permit the tracing of international money flows and the comparison of transactions. As in any other areas, they need to ensure the data meets the needs of those who would use it. Improving the interoperability of data improves the potential for groups to examine the data and identify corrupt practices.

#### **Estimating the benefits of open data**

To estimate the benefits of open data in reducing corruption, we begin with estimates of the costs of corruption in developed countries. There are a few studies that have used cross-jurisdictional data on corruption and growth rates to estimate the impact of corruption. We assume that the European Commission's estimate of 1% of GDP is reasonable.

We think a reasonable contribution of open data in reducing corruption is in the order of 10%. Given that the cost of corruption is assumed to be 1% of GDP which is around AUD 1.5 trillion, the total value of the benefits of open data in reducing corruption can be conservatively estimated at around AUD 1.5 billion per annum or 0.10% of GDP per annum.

## Annex 5 G20 Agenda Item – Employment

It is difficult to understate the importance of employment to economic prosperity and to wellbeing more generally. The effectiveness of the labour force and its deployment is a key driver of national productivity. However, more significant again is the social importance of employment. Studies have consistently shown that the wellbeing of individuals is significantly influenced by employment conditions.<sup>120</sup> For those who are employed, satisfaction with work has repeatedly been found to be strongly correlated with life satisfaction (Judge and Klinger 2008).<sup>121</sup> Research has linked job satisfaction to many other factors of interest including higher job participation (including delaying retirement) and workplace productivity.

### Searching and matching

The matching of workers to workplaces matters; it is important both for the satisfaction of the workers and the productivity to the employer. For example, Binder and Coad (2010)<sup>122</sup> find that individuals moving from regular employment into self-employment experience an increase in life satisfaction (up to two years later). Further, the significance of employee matching for productivity is demonstrated in a recent study by Jackson (2013).<sup>123</sup> Jackson examined the importance of the match between teachers and schools for student achievement. He looked at the difference between better job matching and teacher effectiveness and found that where teachers moved jobs and their match with the work improved, teacher effectiveness was improved and that such changes were as important as teacher quality.

### The problem of matching

Contrast the challenges faced by a job-seeker deciding between alternative places of employment and that of a buyer of a simple consumer product, such as a washing machine. The job-seeker's decision is, of course, far more significant. The impacts, both in terms of employee well-being and work-place productivity, are significant and the costs of switching are high.

Despite its significance, the information available to the job-seeker is relatively limited. For the washing machine, the buyer may have access to substantial information that may help with the purchase decision such as on-line reviews, first-hand experience with other products from the same brand or store and recommendations from friends. In contrast, the job-seeker's information on the prospective work-place will be generally limited to information provided by the firm itself.

### The role of governments and open data

Open data has potential to transform labour markets through transmission of information on work-place conditions to prospective employees. Consider the impact of higher quality information on employment conditions being released to prospective job-seekers. The immediate result would be a better match of worker to organization with benefits of greater job-satisfaction and improved productivity. A longer term benefit would

---

120 A common research finding is that the unemployed consistently report much lower levels of wellbeing than the employed and that difference is far greater than can be accounted for due to the loss of financial resources. Further unemployment is not just associated with losses in wellbeing. It *causes* them. Lower wellbeing may contribute to unemployment, research has confirmed that unemployment leads to lower wellbeing. For example, panel regression analysis consistently finds that wellbeing rises when the unemployed find a job.

121 Judge, T. A., & Klinger, R. (2008). Job satisfaction. *The Science of Subjective Well-Being*, 393 (Eid, M., & Larsen, R. J. Eds.).

122 Binder, M., & Coad, A. (2013). Life satisfaction and self-employment: a matching approach. *Small Business Economics*, 40(4), 1009-1033.

123 Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics*.

be the impact on workplace quality as firms, more concerned than ever with their reputation as employers, seek to improve workplace conditions. This, in turn, should also encourage greater work-force participation.

There have been some private attempts to fill the information gap. Many large firms organize independent surveys of employees to assess job-satisfaction, which, if positive, are released to the market. The organization *Glassdoor* has built a database on companies based on anonymous information supplied by employees. However, the potential for private competitive solutions to the matching problem appear limited. Firms have an interest in only releasing positive information and even here there is no established general standard that would produce results that would allow people to compare individual firms, or even to understand where they fit more generally amongst their peers in terms of economy wide or industry averages.

The issue is similar to a recognised problem in some product markets whereby firms hide (or choose not to disclose) information on some product attributes – such as high charges for optional extras. Gabaix and Laibson (2005) describe such attributes as 'shrouded attributes', attributes 'hidden by a firm, even though the attribute could be nearly costlessly revealed'.<sup>124</sup>

Governments cannot wave a wand and banish such asymmetric information problems. But they already have some tools with which they can begin to make progress. Thus the Australian Public Service Commission (APSC) collects the data from an elaborate questionnaire circulated annually amongst Australia's public servants. It publishes the data but is assiduous in doing so in such a way that inter-agency comparisons cannot be made. In Gabaix and Laibson's terminology, their performance in generating satisfying jobs remains a 'shrouded attribute'! The Commission should publish the data in much more disaggregated form, indeed disaggregating as far as possible and stopping only at the point that compromises the privacy and confidentiality of the answers to those giving them. There are also service providers ready and willing to add value to APSC jobs data (see below).

Further, as documented in the "Windows on Workplaces" concept, governments can help play a role in the emergence of standards. Once such standards exist, it becomes in the interests of the best performing firms to collect and publish data according to that standard. And once that occurs, just as better-informed prospective employees place pressure on managements to improve, so firms that do not publish against the standard leave themselves open to the suspicion that they are not reporting because they are performing poorly. Again the pressure is towards greater and better reporting.

### Quantifying the benefits

As discussed above, the benefits of open data on employment conditions may stem from:

- increased productivity through better employee matching;
- increased employee satisfaction and increased labour force participation; and
- second round effects of workplaces that are better managed as a result of employees being choosier given the additional information available to them.

All these sources of benefit could be very significant. An indication of their magnitude in just one area – teaching – may be inferred by building on the research on the example of teacher quality. There is a large body of research (Jensen 2010 provides a brief summary)<sup>125</sup> that highlights the value of increasing teacher

---

124 Gabaix, Xavier and Laibson, David. 2005. *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*. MIT Department of Economics Working Paper No. 05-18.

125 Jensen, B. (2010). Investing in our teachers, investing in our economy. Grattan Institute.

effectiveness. For example, Hanushek (2011)<sup>126</sup> estimates for the United States that a teacher one standard deviation above the mean effectiveness annually generates marginal gains of over USD 400,000 in present value of student future earnings. He further concludes that replacing 'the bottom 5-8% of teachers with average teachers could move the US near the top of international maths and science rankings with a present value of USD 100 trillion.'

Jensen (2010) leverages the international research to estimate the benefits for Australia. He estimates (page 19) that a 10% increase in teacher effectiveness would lead to 'an increase in the long run GDP growth rate by 0.2% every year, adding AUD 90 billion to Australia's GDP by 2050, and make Australians 12% richer by the end of the century.' However, as noted above, the research by Jackson (2013) indicates improving the match quality (i.e. the quality of matching of the teachers to schools) can have similar effects as increasing the teacher value-added.<sup>127</sup>

Arguably the improving match quality through open data is more easily achieved than improving teacher quality. In this regard, it is notable that being major funders of the vast majority of educational establishments in Australia, governments have strong purchase over educational management and practice and are in a strong position to insist on the generation and release of much better data on job quality within educational establishments, particularly for the benefit of teachers and prospective teachers seeking to find their preferred employer.

Given this, and on the basis of Jackson's results, applied to Australia using Jensen's methodology, we think it very conservative to suggest that such promotion of open data could have the equivalent effect of a 2.5% improvement in teacher effectiveness (i.e. a quarter of Jensen's (2010) aspiration of a 10% increase). On this basis, the potential benefits of open data from improved teacher effectiveness alone would be 0.05% GDP growth. Over the five years contemplated in the G20 target, this would produce gains of close to AUD 800 million in the first year rising to over AUD 4 billion in the fifth year at a present value average over the period of AUD 2.3 billion per year.

But this is just the tip of the iceberg. Productivity benefits from improved job matching could be similar in other areas such as health, community services and government administration, where governments either have control or substantial influence. And where governments have little direct purchase on domains of employment, programs like 'Windows on Workplaces' could create 'lighthouse' effects whereby the best employers could demonstrate their prowess as employers with commensurate matching improvements in their own workforces while pressure would be placed on all other firms both to collect and publish similar data and, as a result of that transparency, to lift their game. A true, which is to say informed, market for job quality and satisfaction would have been established with immense social and economic (productivity) benefits.

And changes of this magnitude would generate substantial indirect benefits. Just one such likely source of benefit would be a reduction in workplace stress, the productivity cost of which has been estimated in Australia to be (in 2008) in excess of AUD 14 billion per year (around 16 billion per year when inflated to the present time).<sup>128</sup> We would expect open data on workplaces would significantly contribute to a reduction in workplace stress through better job matching and through reputational pressure on workplaces to reduce

---

126 Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, 30(3), 466-479.

127 Jackson, C. K. (2013). Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics*. Specifically Jackson (page 3) finds evidence that 'one standard deviation increase in match quality [has] about the same effect as a one standard deviation increase in teacher value-added [on math scores and reading scores].'

128 See Medibank Private, The cost of workplace stress in Australia. (at <http://www.medibank.com.au/Client/Documents/Pdfs/The-Cost-of-Workplace-Stress.pdf>) based on commissioned research by Econtech

workplace stress. A 5-10% improvement appears a reasonable, perhaps conservative, estimate of the reduction in workplace stress that would be achieved through greater transparency on workplace conditions. If so, this contributes between AUD 0.8 and AUD 1.8 billion per annum. The mid-point of these amounts (i.e. AUD 1.2 billion) combined with the estimate from the improvement in teacher effectiveness give a combined total (with adjustment for rounding) of AUD 3.4 billion per annum.

Notably, these estimates do not include the improvements to wellbeing that would flow directly to job-seekers themselves and the greater levels of participation in the workforces.

### **G20 Agenda Item – Employment: The APS example**

Since 1902, the Australian Public Service (APS) Gazette has noted the appointments, terminations, promotions and movements in the Federal bureaucracy, an important source of information for senior public servants and Canberra watchers. More importantly, the public notification of APS employment decisions reinforces the openness, transparency and accountability of the APS.

In 2007, the Australian Public Service Commission started making the Gazette notices available in PDF format and on their website and, as part of the 2013 annual GovHack event (a government sponsored 'mashup' event at which volunteers develop government data to prototype stage), Felix Barbalet (Managing Director of Pivotal Analytics - a Canberra Analytics consulting firm) used the PDF files to create an open and searchable interface to the data.

The major breakthrough was the conversion of the Gazette archives from a pseudo-print PDF format (which does not lend itself to any form of analytics) – to a machine-readable open-data format. This meant it was possible to match a particular employment decision to the original job notice, unlike the APSC website and published Gazette notices where there was almost no information about the original job published along with the hiring decision.

The result was a gold-mine of data for human resources managers within the public service and for private-sector recruiters and hiring managers looking for experienced professionals. Early adopters of the service found huge value in the data because it connected them with experienced potential candidates with a proven track record, many of whom would have been very difficult to access any other way.

### **Quantifying the value to private industry**

Seek – Australia's largest job website – reported revenue from its domestic employment arm of AUD 234.8 million in the 2013 financial year. They reported an average of 150,000 job listings per month; calculating the "value" of a job notice based on their annual revenue and the annual number of job listings, we get an estimated annual value of more than AUD 1.5m for just the job notices portion of the Gazettes in one year. Another key source of value contained in the Gazette data related to the candidates. Recruiters value access to high quality candidates and with the fee for successful placement typically in the range of 10-30% of candidates' annual salary, access to skilled candidates is highly sought after.

Using a service like LinkedIn Corporate to contact candidates costs between AUD 15 and AUD 20 per candidate, and advertising on a major job board results in a cost of around AUD 10 per resume received.<sup>129</sup> Feedback from early adopters of the APSJobs service built by Pivotal Analytics reported that the quality of candidates was higher than those available elsewhere suggesting values above these.

Taking a very conservative estimate of AUD 15 per skilled candidate gives an additional annual value of approximately AUD 600,000 annually for the details of candidates and the positions they've been placed into (based on an annual listings of around 40,000 candidates and hiring managers). It is likely that this is a small fraction of the value generated from better job matching that the site permits.

---

<sup>129</sup> Private communication from a recruiter.

### Quantifying the value to the citizen

Open data encourages innovation and re-use. Another way to quantify the value of open-data to the public is to examine their engagement with sites created using it compared with the sites using the same data created by the public service or sites using similar data from private industry.

**Table 6 APSJobs.Info, APSC and Seek's site performance**

	<i>Bounce rate *</i>	<i>Average page views</i>
APSJobs.info, APSIndex.com (Pivotal Analytics)	8.4%	8.00
APSJobs.gov.au (APSC)	23.5%	2.90
Seek.com	21.1%	8.89

\* Bounce rate is the rate at which visitors to the site leave immediately (lower is better). Average pageviews shows the level of engagement for those visitors who do not leave immediately (higher is better).

This data illustrates the value of open-data to citizens and other users of digital content (for instance business) – it enables the creation of tailored and innovative content that they want to view.

### Obstacles to innovation

One of the key obstacles to continued use of the APS Gazette data was that the APSC have never made it available in a machine-readable format. It is still published weekly in a pseudo-print PDF format, but constant changes to the formatting and layout of these files has meant that continuous re-development of the PDF conversion software is required for ongoing access to the data.

The skills required to perform such work (data-mining and text-analytics) are in great demand, and as such it was not viable for Pivotal Analytics to continue devoting resources to re-develop the conversion software every few weeks. This illustrates why it is so important for government to explicitly embrace open-data and to ensure that government agencies deliver on higher level commitments to open data; otherwise, it is unlikely many private enterprises or citizens will be willing to invest resources in building innovative systems that rely on data sources that may disappear from week to week.

### Future access to the APS Gazette notices

In the wake of privacy concerns expressed by Australia's Parliamentary Joint Committee on Human Rights the APSC is currently conducting a review of the publication of the Gazette notices due to privacy concerns, with the head of the Australian Public Service Commission noting that "Today, termination information published online in the Gazette is available years after the fact, and is discoverable via a search engine".<sup>131</sup>

### Annex 6 G20 Agenda Item – Energy

The energy sector, which incorporates the extraction of energy sources and the supply of energy, is a significant sector in all G20 economies. In Australia – a large energy exporter – this is particularly the case as energy's direct contribution to gross value added is estimated to be around 6% per annum.<sup>131</sup> Small improvements in efficiency can thus contribute to significant impacts.

The potential for open data to drive efficiency improvements in the energy sector has been highlighted by a number of commentators (e.g. McKinsey Global Institute, Gee et al). Key opportunities identified have included efficiencies in:

130 <http://www.smh.com.au/national/public-service/aps-gazette-employment-pages-in-doubt-20140407-367xf.html>

131 BREE - Energy in Australia 2013. At <http://www.bree.gov.au/publications/energy-australia>.

- Resource extraction through, for example, sharing of geological and seismic data to improve the efficiency of exploration activity.
- Supply through better use of data in optimising investment decisions and operations in generation, transmission and distribution.
- Demand (particularly with regard to electricity) through greater use of open data to inform consumption decisions and use by retailers in demand management.

The estimates of potential gains have been large; McKinsey estimates the potential value of open data in the electricity, oil, and gas sectors is USD 11 billion (USD 7 billion for electricity, USD 4 billion for oil and gas).

Australia has made some progress on many of the key issues identified in the McKinsey report. For example, there are programs to provide consumers with information on opportunities to reduce energy consumption<sup>132</sup> and develop and release geospatial information to drive more efficient exploration and development. Furthermore, there appears to be significant barriers in place — for example, rigid pricing structures that reduce the incentives for efficient energy use — that prevent some opportunities being realised.

Nevertheless, there still appear to be significant opportunities. For example, a recent scoping study on consumer energy data access system (CEdata)<sup>133</sup> identified shortcomings with access to existing information. Opportunities will also continue to exist in other areas noted above relating to improving investment decisions and operations management in energy supply and some areas related to demand (an example relates to energy efficiency of buildings, see Box 15 below). Furthermore, there are additional benefits related to other sectors. For example, a benefit of greater use of open data in the transport sector will be a reduction in fuel use stemming from better route selection, more ride sharing and reduced congestion.

A potential benefit of open data is in providing a catalyst for other reforms such as pricing reforms and modification/removal of other programs. For example, greater transparency of the cost burden associated with inefficient programs could encourage consumer demand for change.

### **Box 15 Open data and the energy efficiency of buildings**

#### **Building Performance Institute of Europe (BPIE) Open Data Hub**

Buildings account for 40% of the EU's total energy consumption and 36% of greenhouse gas emissions, and 75% were built during periods with either no or minimal energy efficiency building codes. Accordingly this represents a major opportunity to reduce energy costs and emission.

The BPIE's "Europe's Buildings Under the Microscope" report (2011) projected net lifetime savings (energy saving less investment) of ~EUR 450 billion through to 2050. Key to this effort is the development of open data hub cataloguing the EU building stocks in order to provide public and private sectors a common reference point to develop appropriate strategies.

132 Some energy retailers also provide information on relative energy use; for example, Energy Australia operates an eWise portal that enables customers to track and compare their energy use.

133 Simon Orme, James Swansson and Eli Hefter, (2012). 'Scoping study for a consumer energy data access system (CEdata)', Report for Department of Resources, Energy and Tourism. At <http://www.innovation.gov.au/energy/Documents/energyMarket/CEdata-scoping-study.pdf>.



The Open Data Portal covers: 29 European countries, 10 building types, up to 18 climatic zones per country, a total building stock floor area nearly equivalent to the size of Belgium, and provides data and statistics on: Building stock performance (energy consumption, envelope performance, energy sources), Building stock inventories reflecting floor area, construction year, ownership profile, national policies and regulation, and financial schemes (~333 schemes).

In light of these considerations, we feel it reasonable to assume that open data could at least contribute to conservatively a 2% increase in efficiency of the sector which equates to annual benefit in terms of contribution to GDP of around AUD 1.7 billion per year or 0.11 per cent of GDP.<sup>134</sup>

### **Annex 7 G20 Agenda Item – Infrastructure**

Infrastructure covers a wide range of asset types across the Australian economy. At the national level, the Infrastructure Australia Act 2008, defines "nationally significant infrastructure" to include transport, energy, communications and water. At a local level, it includes playgrounds, parks and roads.

Collectively these assets total almost AUD 2 trillion. In terms of activity, Australia's infrastructure industries contributed 10% of GDP for the year 2011-12, with transport contributing almost half of this. These amounts understate the importance of infrastructure to the national economy as the value derived from infrastructure is far in excess of the cost to develop. Furthermore, infrastructure costs tend to rise through time – as congestion mounts and as the best locations are used for infrastructure, further investments tend to be more expensive per unit of output.

### **Box 16 Examples of better utilisation of infrastructure data**

NICTA offered a number of good examples of how data can drive economic benefits across the infrastructure lifecycle.

*Port Botany Rail:* using data to optimise rail supply chain operations improves throughput on existing infrastructure, delaying AUD 200m track upgrades by 15-20 years.

*WA Roads:* optimisation of traffic flows using smart sensors to increase throughput on existing infrastructure and avoiding AUD 300 millions of roadworks.

Source: NICTA submission to the Productivity Commission inquiry on infrastructure <sup>135</sup>

134 The 1% efficiency increase is applied on the estimate of gross value added of the sector of AUD 79 billion in 2010-11. The estimated benefit is equivalent to a little less than one half of the benefit reflected in the McKinsey report using the approach discussed in section 3.3.2. We felt that, while there are large gains to be had from better data management in this sector, many of them relate to better private data handling – for instance the installation of smart meters. As a consequence we felt more comfortable being a lot more conservative than McKinsey in this area, though we expect our estimate is an underestimate.

135 NICTA submission - <http://www.pc.gov.au/projects/inquiry/infrastructure/submissions#post-draft>

The very high cost of developing new infrastructure to meet growing demands leads to two key infrastructure challenges:

- maximising the value obtained from existing infrastructure, and
- ensuring the efficient development of new infrastructure.

Open data can assist in both areas.

### **Maximising the value of existing infrastructure**

National infrastructure — and in particular transport infrastructure — suffers from issues of under-utilisation during some periods and excess demand leading to congestion in other periods. Improved information made available as part of an open data focus can help improve utilisation in a number of ways.

Information on infrastructure utilisation can help users to improve their use, for example, alerting them to where and when infrastructure is congested. This information may include data on current use or forecasts of future use based on current and historical information. For example, we might expect that data on current and historical use of road-networks will result in organizations developing predictive models to help users better plan their travel.

Improved information may also help other parties, such as those providing services associated with the infrastructure. For example, knowledge on road congestion may be used by a port operator for planning purposes. A further benefit of improved utilisation information is that it may provide a foundation to facilitate other regulatory reform (e.g. associated with congestion pricing). The benefits from better utilisation are two-fold. Better utilisation of the infrastructure can increase the value of existing infrastructure (see roadworks example); and result in postponement of building new infrastructure (see examples from NICTA).

### **Box 17 Australia's AUD 10.5 billion local government public infrastructure opportunity**

Australia has 562 local government authorities that manage over AUD 210 billion of public infrastructure assets. The consumption of these assets represents an annual cost of AUD 6.3 billion, based on an average annual deterioration rate of 3%.

Based on studies demonstrating 1.5-2.5% reduction in deterioration rates can be achieved through more comprehensive approaches to asset management, a modest 1% improvement across the local government sector has the potential to deliver AUD 10.5 billion worth of savings over five years.

However, while significant reductions in consumption costs can be achieved at a local authority level, an open data approach to infrastructure and asset information will enable this benefit to be realised more rapidly through sharing of best practice information, better industry benchmarking and facilitating the harmonisation of asset data. In addition, open data offer wider benefits accruing from exposing assets to cross authority, state-local, local-national and public-private asset procurement, re-use and renewal opportunities.

### Ensuring efficient development of new infrastructure

Since 2008, the annual spend of new infrastructure has been estimated as more than 2% of GDP.<sup>136</sup>

As noted by the Productivity Commission, ensuring that this spend is efficient is a key challenge.

Greater transparency and greater capture and release of data may play a key role in a number of ways.

- *Improve scrutiny of key infrastructure decisions.* Greater transparency around infrastructure decision (including the assessments of cost and benefits) can improve the scrutiny of infrastructure build and minimise the risk of poor decisions.
- *Improve analysis of costs and benefits.* Improved data can lead to better analysis of demand and the potential to support infrastructure decisions.
- *Improve implementation planning and coordination.* Improved transparency of planning can lead to more optimised roll-out of infrastructure and more efficient use of resources.
- *Improved data can aid procurement process* (see Box 18)

#### Box 18 Efficiency of procurement

As US and UK public sector organizations have opened their data sets, Spikes Cavill have developed on-line products and services to create economic value through the capture, publishing and benchmarking of contract, procurement and expenditure data.

In the UK they work with local authorities to publish spending information to provide transparency for their communities, as well as to identify opportunities such as collaborative procurement and asset re-use.

In the US one customer, Washington Department of Transport, saved AUD 32 million over two years through cross agency collaboration enabled by open data, including AUD 12 million in contract administration saving and AUD 12 million in procurement transaction cost efficiencies.

The opportunity to improve data transparency appears large. The draft Productivity Commission report into infrastructure highlights a number of issues. The inquiry found that:

*"Data problems beset the detailed analysis of the costs and productivity of public infrastructure construction, and of the effects of various policies. A coordinated and coherent data collection process can address this and improve future project selection decisions."*

As noted in the report, current public sector infrastructure asset information is fragmented across both industry and government sectors. Local, state and federal government maintain disparate asset data. At a sector level, approaches to develop and maintain infrastructure vary. The fragmentation leads to inefficiencies from cross sector overlap and misalignment. This occurs due to differing approaches to the provision of infrastructure across sector and government jurisdictions that give rise to resource allocation distortion and inefficiencies where substitution effects are not properly considered. Exposing current "business-as-usual" asset data would represent a significant first step in understanding gaps and begin to provide new evidence-based insights.

136 See page 5 of Productivity Commission 2014, Public Infrastructure, Draft Inquiry Report, Canberra. [http://www.pc.gov.au/\\_data/assets/pdf\\_file/0007/134674/infrastructure-draft-volume1.pdf](http://www.pc.gov.au/_data/assets/pdf_file/0007/134674/infrastructure-draft-volume1.pdf)

### Box 19 Examples – UK roadworks open data

Launched in May 2012, [roadworks.org](http://roadworks.org) is the UK's national road works database, using data from over 140 different local highway authorities.

Part of an industry drive to save costs and maximise efficiency by promoting partnership working across highway authority boundaries, its purpose is to reduce journey times on the UK network by providing reliable, instant and up-to-date information about road works.

It currently publishes the details of 7,000 new road works started in Britain each working day. This data is accessed by motorists, the emergency services, businesses, local authorities and utility providers. The benefit cost ratio of the investment in roadworks has been estimated as 9:1 for councils and 36:1 taking into account wider benefits to the public.

Source: [roadworks.org](http://roadworks.org)

### Estimating the benefits

The potential benefits of open data in improving infrastructure use are significant. Given the current challenges and the potential opportunities described above, it appears reasonable that an open data focus could lead to a 5% improvement in the efficiency of new infrastructure expenditure. Based on current infrastructure spend, the potential benefit is around 0.10 per cent of GDP or a little over AUD 1.5 billion per annum.

This does not include the benefit of improved utilisation of existing assets. The benefits of improved utilisation are likely to largely stem from transport infrastructure where congestion issues are most significant. A useful starting point is the avoidable costs of congestion. These are forecast to reach AUD 20 billion per year in Australia by the year 2020. Through better planning and shifting of traffic, an open data focus can help to reduce a significant portion of this cost. We assume a 10% reduction (AUD 2 billion) in this cost is a reasonable estimate of the potential benefit. Combined, the two sources of benefit are equal to AUD 3.6 billion per year (including an adjustment for rounding).