



Australian Government
Productivity Commission

Data Availability and Use

Productivity Commission Issues Paper

April 2016

The Commission has released this issues paper to assist individuals and organisations to prepare submissions.

It contains and outlines:

- the scope of the inquiry
- the Commission's procedures
- matters about which the Commission is seeking comment and information
- how to make a submission.

The Issues Paper

The Commission has released this issues paper to assist individuals and organisations to prepare submissions to the inquiry. It contains and outlines:

- the scope of the inquiry
- the Commission's procedures
- matters about which the Commission is seeking comment and information
- how to make a submission.

Participants should not feel that they are restricted to comment only on matters raised in the issues paper. The Commission wishes to receive information and comment on issues which participants consider relevant to the inquiry's terms of reference.

Key inquiry dates

Receipt of terms of reference	21 March 2016
Initial submissions due	29 July 2016
Release of draft report	Early November 2016
Draft report public consultation	End of November 2016
Post-draft submissions due	19 December 2016
Final report to Government	21 March 2017

Submissions can be lodged

Online:	www.pc.gov.au/inquiries/current/data-access
By post:	Data Availability and Use Productivity Commission GPO Box 1428 Canberra ACT 2601

Contacts

Administrative matters:	Tracey Horsfall	Ph: 02 6240 3261
Inquiry content:	Rosalyn Bell	Ph: 02 6240 3308
Freecall number for regional areas:	1800 020 083	
Website	www.pc.gov.au	

The Productivity Commission

The Productivity Commission is the Australian Government's independent research and advisory body on a range of economic, social and environmental issues affecting the welfare of Australians. Its role, expressed most simply, is to help governments make better policies, in the long term interest of the Australian community.

The Commission's independence is underpinned by an Act of Parliament. Its processes and outputs are open to public scrutiny and are driven by concern for the wellbeing of the community as a whole.

Further information on the Productivity Commission can be obtained from the Commission's website (www.pc.gov.au).

Terms of reference

PRODUCTIVITY COMMISSION INQUIRY INTO DATA AVAILABILITY AND USE

Productivity Commission Act 1998

I, Scott Morrison, Treasurer, pursuant to Parts 2 and 3 of the *Productivity Commission Act 1998*, hereby request that the Productivity Commission (Commission) undertake an inquiry into the benefits and costs of options for increasing availability of and improving the use of public and private sector data by individuals and organisations.

Background

The 2014 Financial System Inquiry (the Murray Inquiry) recommended that the Government task the Commission to review the benefits and costs of increasing the availability and improving the use of data. The 2015 Harper Review of Competition Policy recommended that the Government consider ways to improve individuals' ability to access their own data to inform consumer choices. The Government has agreed to pursue these two recommendations.

The Australian Government seeks to consider policies to increase availability and use of data to boost innovation and competition in Australia and the relative benefits and costs of each option.

Effective use of data is increasingly integral to the efficient functioning of the economy. Improved availability of reliable data, combined with the tools to use it, is creating new economic opportunities. Increasing availability of data can facilitate development of new products and services, enhance consumer and business outcomes, better inform decision making and policy development, and facilitate greater efficiency and innovation in the economy. As in Australia, international governments are encouraging greater use of data through open data policies. This will increase the transparency and accountability of government processes.

Increased sharing of data across the public and private sectors could facilitate greater leveraging of technology to improve individuals' and entities' interactions with government, improve the integrity of systems and increase administrative efficiency.

In taking advantage of greater use of data, it is important to give appropriate attention to other interests such as privacy, security and intellectual property.

Scope of the inquiry

The Commission is to conduct a broad ranging investigation into the benefits and costs of options for improving availability and use of data. In developing recommendations, the Commission is to:

-
1. Examine the benefits and costs of options for increasing availability of public sector data to other public sector agencies (including between the different levels of government), the private sector, research sector, academics and the community. Where there are clear benefits, recommend ways to increase and improve data linking and availability. The Commission should:
 - a) identify the characteristics and provide examples of public sector datasets that would provide high value to the public sector, research sector, academics and the community to assist public sector agencies to identify their most valuable data; and
 - b) examine legislation or other impediments that may unnecessarily restrict the availability and linking of data, including where the costs are substantial, and consider options to reduce or remove those impediments.
 2. Examine the benefits and costs of options for increasing availability of private sector data for other private sector firms, the public sector, the research sector, academics and the community. Where there are clear benefits, consider ways to increase and improve availability. The Commission should:
 - a) identify the characteristics and provide examples of private sector datasets that would provide high value to the private sector, public sector, the research sector, academics and the community in developing or providing products and services, undertaking research and developing policy;
 - b) identify the concerns of private sector data owners and provide recommendations on principles or protocols to manage these concerns;
 - c) examine legislation or other impediments that unnecessarily restrict the availability of data, including where the costs are substantial, and consider options to reduce or remove those impediments; and
 - d) provide an update on existing data sharing initiatives in Australia, including the uptake of the credit reporting framework. Consider recommendations for improving participation in such initiatives.
 3. Identify options to improve individuals' access to public and private sector data about themselves and examine the benefits and costs of those options. The Commission should:
 - a) examine how individuals can currently access their data, including data about them held by multiple government agencies, and develop recommendations to streamline access;
 - b) identify datasets, including datasets of aggregated data on consumer outcomes at the product or provider level, that would provide high value to consumers in making informed decisions and any impediments to their use. Develop guidance to assist in identification of other high value datasets; and
 - c) examine the possible role of third party intermediaries to assist consumers in making use of their data.
 4. Examine the options for, and benefits and costs of, standardising the collection, sharing and release of public and private sector data.

-
5. Examine ways to enhance and maintain individuals' and businesses' confidence and trust in the way data are used. Having regard to current legislation and practice, advise on the need for further protocols to facilitate disclosure and use of data about individuals and businesses while protecting privacy and commercial interests and, if recommended, advise on what these should be. The Commission should:
- a) balance the benefits of greater disclosure and use of data with protecting the privacy of the individual and providing sufficient control to individuals as to who has their information and how it can be used;
 - b) benchmark Australia's data protection laws, privacy principles and protocols against leading jurisdictions;
 - c) examine whether there is adequate understanding across government about what data can be made openly available given existing legislation;
 - d) consider the effectiveness and impacts of existing approaches to confidentialisation and data security in facilitating data sharing and linking while protecting privacy; and
 - e) consider the merits of codifying the treatment and classification of business data.

In developing its recommendations, the Commission should take into account the Government's policy to improve the availability and use of public sector data (the Public Data Policy Statement) as part of its National Innovation and Science Agenda and to improve government performance through the Efficiency through Contestability Programme, as well as the findings of the Public Sector Data Management Project.

The Commission should consider domestic and international best practice and the measures adopted internationally to encourage sharing and linking of both public and private data.

Process

The Commission is to undertake an appropriate public consultation process, inviting public submissions and releasing a draft report to the public. A final report should be provided to the Government within 12 months from the date of receipt of the reference.

S. MORRISON

Treasurer

[Received 21 March 2016]

Contents

Terms of reference	iii
1 What this inquiry is about	2
2 Benefits of increasing data availability and use	9
3 Managing the costs	21
References	29
Attachment A — How to make a submission	32

1 What this inquiry is about

What has the Commission been asked to do?

The Australian Government has requested the Commission to conduct a broad ranging investigation into the benefits and costs of increasing the availability and use of public and private data by Australian individuals and organisations, including individuals' access to data about themselves. The Commission has been asked to provide recommendations on increasing data access while giving appropriate attention to ways to enhance individuals' and businesses' confidence and trust in the way data is collected, stored and used, with particular focus on privacy and data security.

The inquiry has its origins in the 2014 Financial System Inquiry (the Murray Inquiry) (the inquiry's final report and the Government's response were released in December 2014 and October 2015 respectively) and the Harper Review of Competition Policy (the review's final report and the Government's response were released in March 2015 and November 2015 respectively).

What is data?

A distinction should be made between 'data' and 'information'. Data comprises raw, unorganised material such as characters, text, words, numbers, pictures, sound or video. However, without being put into context, it has little if any meaning. Information is derived from data after it has been processed and presented in context — in this way, data is given meaning.

Some of the main types of data are defined in box 1 and the relationships between them are illustrated in figure 1.

Box 1 **Definitions**

Big data is an overarching term that describes any voluminous amount of data that has the potential to be mined for information. While people and societies have long gathered and stored large amounts of data for eventual analysis, the concept of ‘big data’ is relatively new. Big data has been characterised by the so-called ‘three Vs’:

- high volume — refers to the sheer volume of data being collected
- high velocity — refers to the great speed at which data is being generated, often in near-real time, and how it can be rapidly accessed, processed and analysed
- high variety — refers to the many different formats of data and its diverse sources.

Within big data, there is structured, semi-structured and unstructured data. **Structured data** is characterised by a high degree of organisation. It is commonly text files, displayed in titled columns and rows that can easily be ordered and processed by data mining tools and readily searchable by simple search operations. It largely comprises data contained in databases and spreadsheets. **Unstructured data** is data that is not organised in this way — such as email messages, word processing files, PDF files, digital images, video, audio and social media posts. **Semi-structured data** is a cross between structured and unstructured data — for example, emails have the sender, recipient, date, time and other fixed fields added to the unstructured data of the email message.

Whether data is **internal** or **external** depends on the perspective of the data holder. Data created within a business, government agency or other organisation is internal data, while data sourced from outside the organisation is external data.

Public sector data is data collected and stored by government. **Private sector data** is data collected and stored by individuals, businesses and other non-government organisations.

Personal data is data that identifies or could identify someone, such as their name, address, medical records, bank account details, photos, videos, personal preferences, opinions and occupation — essentially, any data by which someone may be reasonably identifiable.

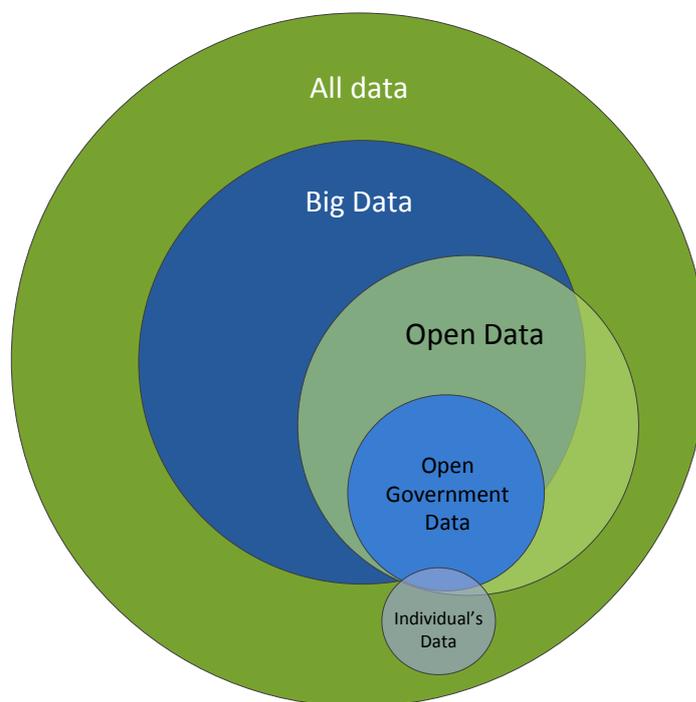
Open data is data that anyone can access, use or share — free or at negligible cost — subject only, at most, to the requirement to attribute and share-alike. Most definitions of open data also specify that it can be retrieved and meaningfully processed by a computer application (machine readable) — that is, it is ‘technically open’ (downloadable in a useful format, for example, an Automated Programming Interface (API) subscription feed). It should also be easy to locate — for example, through an internet search engine. Open data can be internal or external, and collected and stored by the public or private sectors.

Metadata is ‘data about data’. It describes, explains, locates, or otherwise makes it easier to retrieve and use digital resources such as computer files, webpages and databases. For example, the metadata for a digital image may provide details on the size of the image, its resolution and when the image was created.

A distinction is made between **raw data** (including field data) — that is, the basic material that is created and collected — and data that has been ‘cleaned’ or otherwise transformed with the intent to add value — for example, through data mapping and code generation. Raw data can be structured or unstructured.

Sources: Office of the Australian Information Commissioner (2015b); PricewaterhouseCoopers (2014).

Figure 1 Relationships between types of data



Source: McKinsey Global Institute (2013).

Why does data matter?

The amount of data being generated, collected and stored has been growing exponentially. Data is generated from a multitude of transactions, production activities and communications via information and communication technologies, including the internet, and billions of sensors that are capturing activities in the physical world (OECD 2015a).

By some estimates, the amount of data generated worldwide in 2002 (five billion gigabytes) is now being generated every two days, while other estimates suggest that 90 per cent of the world's information was generated in just the past two years (IBM 2016). Since it is unlikely that governments have substantially increased investment in data activity in such a short period, it is reasonable to assume that most of this growth is in private (or personal) data.

Data is increasingly integral to how economies function. Increases in computing power, data storage, data capture and matching technologies are increasing the scope for businesses, individuals and governments to make use of large databases. Analysis of large volumes of data is driving improvements in — and the development of new — products, processes, organisational methods and markets. For example, big data analytics is enabling new perspectives to be gained on the human genome, with potential spin-offs for health and medicine.

However, much of the data generated remains under-utilised. According to one estimate, in 2013, around 22 per cent of data generated in the ‘digital universe’ was potentially useful as an input into subsequent analysis but less than 5 per cent of that data was actually analysed (EMC Corporation 2014).

For governments, the pace of change strengthens the need for monitoring, reviewing and adapting policy settings to ensure they do not unnecessarily or inadvertently constrain the capacity of businesses — and public sector service providers — to be innovative in their use of data.

While governments must be mindful of the legitimate privacy concerns of individuals, and how the ‘digital universe’ is enabling detailed profiles of individuals to be built and used, efficient data management requires more than just privacy standards. Responsiveness and flexibility in managing the availability and use of data (public and private sector) to ensure it provides the maximum net benefit to society is needed in the face of the major shifts now evident in the technology for creating and analysing data.

One hardy perennial is the suggestion that governments should substantially improve their own service provision by using data to strengthen the evidence base for their policies and to improve regulatory implementation and enforcement. But developments in data management also appear to offer governments, businesses and innovators the opportunity to reshape markets and alter previously-accepted paradigms of disadvantage and societal need.

This inquiry will allow exploration of those opportunities, drawing on the active participation of parties with knowledge of big data and its ability to induce change.

Why data is a key economic resource

The fact that data can be shared, used and reused an unlimited number of times (it is ‘non-rivalrous’ in consumption) makes it an especially valuable resource. However, because it is possible for a collector, creator or storer of data to exclude others from accessing that data and using it for their own purposes, data cannot be considered a ‘public good’.¹ These excludability characteristics of data provide strong incentives for businesses to collect and add value to data — for instance, to improve the efficiency of their operations and identify new opportunities. It is often in the interests of businesses to exclude access to their data by other parties or, alternatively, to sell their data to other parties or negotiate other mutually beneficial data sharing arrangements.

¹ Public goods are a type of market failure. They are goods or services that can be ‘consumed’ without reducing their availability to others (non-rivalrous) and from which no one can be excluded from consuming (non-excludable). Two commonly cited examples of public goods are national defence and lighthouses.

The terms of reference for this inquiry ask the Commission to provide an update on existing data sharing initiatives in Australia and consider recommendations for improving participation in such initiatives. The public interest in encouraging any such sharing needs to be clearly identified and used to guide policy development.

On the production side, data is both an input into and a product of transactions — including commercial transactions and non-commercial transactions. An individual looking for information on a search engine (a non-commercial transaction) generates data on their interest and often their location. The websites they select from the search then feed into improving the next search for the same information. This self-generating characteristic of some (but not all) data helps to distinguish it from capital and labour inputs to production.

Economies of scale and scope

For many Australian businesses, data could be considered just another resource that can help them run their operations more effectively and raise their awareness of the needs of their customers. For some businesses, increases in the amount of data they collect or otherwise access can enable them to increase the quality of the services they provide. In turn, this can attract more customers and enable the business to generate more data (economies of scale). Similarly, the linking of diverse sources and types of data can generate value greater than the sum of the parts, such as profiling of individuals and the associated potential to improve the targeting of advertising (economies of scope).

Many data-driven services and platforms, such as social networking sites, are characterised by large network effects where the utility of the services increases more than proportionately with the number of users (OECD 2015a). For example, the benefits to an individual member of a networking app such as WhatsApp tend to increase as more family members, friends and other people become members. Network effects can also be present when companies produce products — for example generators and engines — with embedded sensors that provide the original supplier with a stream of performance data.

Data ownership

While data can be subject to copyright and other intellectual property rights, the concept of ownership of data is sometimes not quite as straightforward as the concept of ownership of physical goods, where the owner typically has exclusive rights and control. For example, telephone call metadata — that is, the location, time and duration of phone calls — could be considered as either the property of the individual caller or the relevant phone network.

Stakeholders laying claim to ownership of particular data can include, for example, parties that create the data, parties that commission the data creation, parties that purchase the data and parties that are the subject of the data. In cases where data is considered ‘personal’, ownership rights can be particularly complex because privacy regimes typically tend to

strengthen the ownership rights of individuals over data to which such rules apply (OECD 2015a). For example, do the medical records generated by a patient visiting a doctor belong to the patient, the doctor, or both parties? This inquiry will consider whether ownership of data is essential to achieving privacy.

With personal devices such as smart phones and Fitbits collecting data on the activities of individuals, questions arise as to who owns the data when the personal device communicates with a wider network. There are also questions around what constitutes an individual's 'consent' for an organisation to collect, use and share data about them — and debate over the distinctions between 'informed consent', 'passive consent', 'unknown consent' and 'non-consent'. For instance, does an individual's agreement to Facebook's *Data Use Policy*, which is available to prospective members to read before they agree to join Facebook (and remains available after joining), constitute meaningful consent to the collection, use and sharing of their data? Or, in the alternative case, to what degree should the buyer beware when signing up to an apparently free but nevertheless valuable service?

The Commission's approach and previous work

In preparing its report to Government, the Commission will analyse the costs, benefits and risks of increasing access to data and examine ways to increase data availability and use. To avoid replication, the inquiry will draw on recent studies and reviews — for example, the findings of the Public Sector Data Management Project (Australian Government Department of the Prime Minister and Cabinet 2015). It will also take into account the Government's recent policy announcements to improve availability and use of public sector data as part of its Innovation Statement, including its Public Data Policy Statement (Australian Government 2015b).

The Commission will conduct its own analysis and draw heavily on input from participants through consultations, written submissions (see attachment A for how to make a submission) and stakeholder comments, as well as any relevant previous Commission work. Box 2 examines some of the main recurring data-related themes to have emerged from the Commission's work in recent years. The inquiry will also draw, where relevant, on work currently underway in the Commission's inquiry on Education Evidence Base. The Commission will seek to ensure that its recommendations to Government are consistent with the best outcomes for the wider community.

The inquiry will draw on overseas experience in public and private sector data availability and use, and evaluate approaches used in overseas jurisdictions, such as the integrated data infrastructure in New Zealand and the Midata program in the United Kingdom. Where possible, the Commission will also draw on learnings from overseas debates, such as issues around individuals' access to, control over and retention of data about them, as well as issues experienced with and lessons learned from data security, data breaches and data sharing.

The Commission will also benchmark Australia’s data protection laws, privacy principles and protocols against leading overseas jurisdictions. The Commission encourages stakeholders to provide information about leading practices overseas.

Box 2 **Recurring data-related themes arising in Commission reports**

Previous Commission reports have cited a number of issues surrounding access and use of public sector data (and to a lesser extent private sector data). Recurring data-related themes include the following:

- **Insufficient data sharing between agencies.** Sharing of data between government departments and agencies does not occur as often as it could. This can result in unnecessarily duplication of regulatory data requirements for businesses and individuals, and also impede the quality of the evidence base for future policy making. See, for example, *Annual Report 2012--13* (PC 2013a), *Caring for Older Australians* (PC 2011a) and *Regulator Engagement with Small Business* (PC 2013b).
- **Insufficient dataset linkage.** Related administrative datasets are often not linked together. This, too, leads to inefficiencies and repeated data collection for businesses and individuals and also stymies potential research. Evidence suggests that health-related research, in particular, could benefit from greater integration (*Efficiency in Health*, PC 2015a). See also *Barriers to Effective Climate Change Adaptation* (PC 2012a), *Public Infrastructure* (PC 2014b) and *Childcare and Early Childhood Learning* (PC 2014a).
- **Little public access to administrative data.** Public sector data collected in the provision of services and regulation of industries is, more often than not, inaccessible to members of the public. Not only does this run counter to notions of government transparency, but it also impedes consumer decision making by preventing the public from obtaining information that could help them make the best choice among service providers. See, for example, *Disability Care and Support* (PC 2011b), *Gambling* (PC 2010c) and *Contribution of the Not-For-Profit Sector* (PC 2010b).
- **Limited data access for research.** It can be difficult for researchers to obtain administrative datasets for use as empirical evidence. Compared to administrative data, surveys can have certain limitations for research purposes because of small sample size, selection bias and respondent drop-out. See, for example, *Housing Assistance and Employment in Australia* (PC 2015b), *Childcare and Early Childhood Learning* (PC 2014a), and *Caring for Older Australians* (PC 2011a).
- **Non-standardised datasets.** The absence of standard formats in which datasets are presented and poor quality metadata also inhibit dataset linkage. See, for example, *Impact of COAG Reforms: Business Regulation and VET* (PC 2012b) and *Education and Training Workforce: Schools Workforce* (PC 2012d).
- **Missed potential for stronger evidence-based policy.** The Commission has identified a number of benefits of evidence-based policy development (PC 2010a). Greater use of administrative datasets could help to assess whether policies and programs meet their stated objectives and operate as intended. The Commission has also identified lack of data as a barrier to improving the quality of government regulation and its enforcement. See, for example, *Annual Report 2012--13* (PC 2013a), *Regulatory Impact Analysis: Benchmarking* (PC 2012c) and *Regulator Engagement with Small Business* (PC 2013b).

Sources: PC (2010a, 2010b, 2010c, 2011a, 2011b, 2012a, 2012b, 2012c, 2012d, 2013a, 2013b, 2014a, 2014b, 2015a, 2015b).

2 Benefits of increasing data availability and use

Increased availability and use of data can promote a range of potential benefits across the private and public sectors, including:

- efficiency — by replacing traditional and intuitive approaches with data-driven processes (for example, to better understand consumer preferences or the effectiveness of public programs), data can either lower the costs that businesses and governments incur in providing goods and services or allow them to better target their products to consumers
- empowerment of consumers — increased access to the data created through everyday transactions can empower consumers to make decisions based on what best suits their situation; similarly, provision of data on the relative offerings and performance of product and service providers can help consumers to assess what is available
- competition — wider availability of data can create market opportunities for new businesses, or enable existing businesses to expand into new areas, thus fostering more competitive markets
- innovation — data can help to break down information gaps within and across parts of the economy, potentially providing the building blocks for new products and processes
- accountability of governments — public sector data can shed light on the effectiveness of existing and past government interventions, improve the design of future policies and programs, enable community scrutiny of the evidence base (such as government-funded research) used to support policy interventions, and generally sharpen incentives for governments to perform well.

These benefits can translate into better outcomes for consumers (lower prices, improved product offerings), higher levels of productivity and economic growth, and improved performance of governments, including through better-informed, more effective policies.

A number of authors have examined the broader societal implications of the digital age — see, for example, Brynjolfsson and McAfee (2014), who conclude that the ‘transformations brought about by digital technology will be profoundly beneficial ones’, albeit accompanied by economic disruption and other ‘difficult challenges and choices’. Numerous studies have attempted to quantify the potential benefits arising from improvements in data availability and use. While some of these have indicated the potential for substantial benefits globally (see, for example, McKinsey Global Institute 2013), others — such as Robert Gordon and Joel Waldfogel (New York Times 2013) — have suggested more modest benefits, noting, for example, that some of the cited benefits come from displacing traditional advertising, media and retailing businesses.

The diversity of views and estimates of impacts is indicative of the inherent uncertainties involved and the wide variety of methodologies used to derive estimates (for a brief overview of some of the studies conducted relating to the potential value of public sector

data see Houghton (2011) and Australian Government Department of Communications and the Arts Bureau of Communications Research (2016)). Further, in deriving net benefits to the community as a whole, account must also be taken of the costs of increasing data availability and use (discussed below).

Public sector data

Australian governments collect, create and produce a wide variety of data including legal, administrative, regulatory, business, economic, geographic and meteorological data. Some of this data, when securely de-identified, constitutes a resource that has the potential for use not just within government but also by businesses, researchers, academics and the broader community. Australian governments make some of their data publicly available — for example, through agencies such as the Australian Bureau of Statistics, the Bureau of Meteorology, Geoscience Australia and PSMA Australia Limited (the G-NAF database (box 3)), as well as through websites such as My School and reports such as the Commission’s annual *Report on Government Services*. In addition, some government institutions provide searchable databases — for example, the National Library of Australia and the Australian War Memorial. While there has been considerable progress by Australian governments towards increasing the amount of open data — particularly geographic and other non-personal datasets — not all public sector data is being utilised to its full potential, particularly in view of the recent rapid growth of computing power and data analytics technologies.

There appears to be scope for improved sharing and use of public sector data within and between governments and with businesses, researchers and individuals. Major public sector datasets in social policy fields are often made inaccessible even to neutral third parties (such as academic researchers) by the various regulatory and other processes surrounding the release of datasets containing personal data on individuals. Further, some public sector datasets contain data provided by businesses to government agencies under contractual arrangements that inhibit data release. Evaluation of program effectiveness and improvements in future policy design are examples of benefits that could result from making key datasets more accessible to third parties.

Box 3 **The Geocoded National Address File (G-NAF)**

What is G-NAF?

G-NAF is a database of all the physical addresses in Australia. Because there is no single organisation that is responsible for assigning and collating addresses in Australia, G-NAF is built from addresses supplied by ten contributors, including the land agencies in each Australian state and territory.

As each contributor collects and stores addresses differently, the G-NAF production process involves independently examining and validating every candidate address followed by a process of textual and spatial matching.

Addresses from different sources found to be identical are merged into a single G-NAF record with feature level metadata capturing its lineage and quality. Through this process over 30 million contributed addresses are distilled into over 13.5 million G-NAF addresses.

The presence of accurate information about addresses can deliver benefits in a broad range of contexts, ranging from commercial developers to individuals needing to pinpoint their location in an emergency. Such data provides a geographic context that can assist decision making and promote innovation based on quality location data.

Recent release

In February 2016, G-NAF was made openly available through a cooperative agreement between PSMA Australia and the Commonwealth and state and territory governments. Media reports immediately following the release suggest that there has already been significant take up of the database by commercial developers.

Source: PSMA Australia (2016).

In addition, providing greater ability for potential users to link different datasets could help generate a range of benefits:

Data linkage can consolidate administrative data with information held elsewhere, such as surveys. Administrative data can indicate what happened to whom in terms of pathways and outcomes benchmarked against policy variations. Surveys can elicit more targeted information on why people behaved as they did. A further benefit of data matching would be to enable surveys to omit sensitive questions, such as income levels, substance abuse or other factors that typically get a low response. This would reduce costs and respondent burden. (PC 2013a, p. 4)

Illustrating the potential value of public sector datasets to users outside of government, the Garvan Institute in Sydney recently signed an agreement with the United Kingdom Department of Health — the United Kingdom being ranked as one of the top three OECD countries for accessibility and sharing of health data (OECD 2015c) — to sequence genomes from 100 000 patients with cancer and rare diseases. The resulting research could identify disease causes, characteristics and, potentially, cures, thus generating substantial community-wide benefits (Garvan Institute 2015).

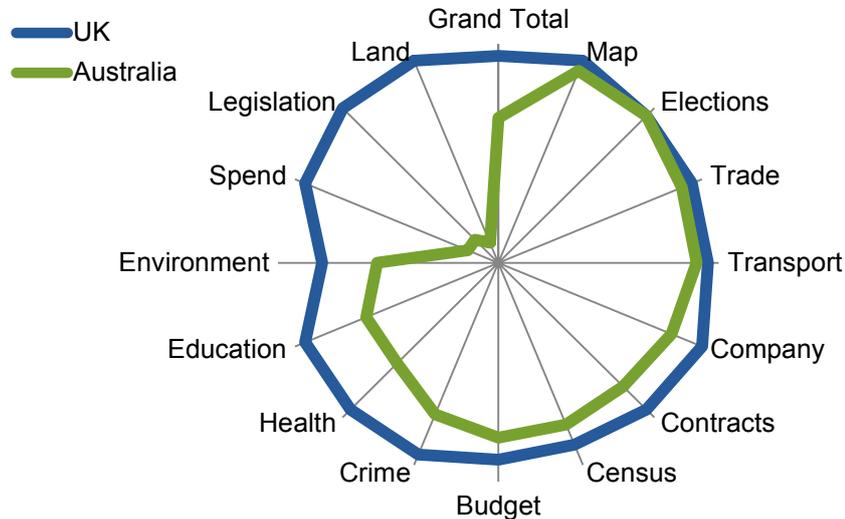
Australia's commitment to public sector open data

Internationally and in Australia, open data policies have progressed significantly over the past five years. There are several sets of international rankings that compare the status of public sector open data across a range of countries based on measures such as availability, accessibility and impact. Australia is currently ranked 4th in the OECD's OURdata Index (OECD 2015b), 5th in the Global Open Data Index (Open Knowledge 2015) and 10th in the World Wide Web Foundation's Open Data Barometer (World Wide Web Foundation 2015). By contrast, the United Kingdom — an international leader in open data — is ranked 3rd, 2nd and 1st respectively. Compared to the United Kingdom, Australia appears to have low levels of availability and accessibility for data relating to the environment, government spending, legislation and land use — although this could reflect, at least in part, disparate data at a sub-national level (figure 2).

The National Commission of Audit (Australian Government 2014) noted that Australian governments had, at that time, only released around 3200 datasets, compared to 10 000 datasets in the United Kingdom and 200 000 in the United States. While it is difficult to assess progress from numbers of datasets alone, the discrepancy is notable. That said, the release of datasets does not automatically generate benefits for the community — the datasets concerned need to be in a usable format and of potential interest to other parties.

The Australian Government's Public Sector Data Management project noted that Australia lags the United Kingdom and the United States in releasing public sector data for business, and New Zealand in regard to applying data to government policy design. It observed that all of these countries had made an upfront commitment to drive data policy with a 'top-down mandate from Ministers'. It also noted that 'sustained action and commitment was key' (Australian Government Department of the Prime Minister and Cabinet 2015, p. 5).

Figure 2 **Comparisons of open data availability^a**
2014



^a Scores for each data type are calculated using a formula that awards points according to a range of factors, such as format and availability.

Source: World Wide Web Foundation (2015).

Following the Australian Government's *Declaration of Open Government* (Tanner 2010), the Office of the Australian Information Commissioner developed a set of principles to guide the availability of open public sector information (public 'open data'). The default position is stated as being 'open access to information' on the basis that 'information held by Australian Government agencies is a valuable national resource' (Office of the Australian Information Commissioner 2011b). This inquiry will seek to determine if that commitment is being met.

More recently, the Australian Government released a *Public Data Policy Statement* in which it committed to:

... optimise the use and reuse of public data; release non sensitive data as open by default; and to collaborate with the private and research sectors to extend the value of public data for the benefit of the Australian public. (Australian Government 2015b, p. 1)

Again, efforts will be made by the Commission to put this commitment to the test.

QUESTIONS ON HIGH VALUE PUBLIC SECTOR DATA

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

What characteristics define high-value datasets?

What benefits would the community derive from increasing the availability and use of public sector data?

QUESTIONS ON COLLECTION AND RELEASE OF PUBLIC SECTOR DATA

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

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

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

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

What specific government initiatives (whether Australian Government, state, territory or local government, or overseas jurisdictions) have been particularly effective in improving data access and use?

QUESTIONS ON DATA LINKAGE

Which datasets, if linked or coordinated across public sector agencies, would be of high value to the community, and how would they be used?

Which rules, regulations or policies create unnecessary or excessive barriers to linking datasets?

How can Australia's government agencies improve their sharing and linking of public sector data? What lessons or examples from overseas should be considered?

Private sector data

As noted earlier, the recent explosive growth in digital data means that large quantities of data are now held by the private sector. Private data holdings present different issues for public policy (indeed, in some cases, they may present no issues at all). The private sector collects and uses many different types of data — for instance: mining companies collect geological data in the course of undertaking exploration; banks collect data on individuals' financial transactions; embedded chip technology is turning former products into services via wireless reporting; and road toll companies, airlines and, increasingly, urban public transport operators collect data on individuals' travel activity.

The extent to which individuals and businesses are prepared to make their data available to other parties will generally be based on their assessment of the likely benefits and costs to them of doing so. For example, individuals implicitly agree to share some personal details every time they complete a transaction using a credit card because of the convenience and other benefits that such cards provide. Similarly, internet users are sometimes given the option of consenting (or not consenting) to tracking cookies — which are commonly used as ways to compile long-term records of individuals’ internet browsing histories — being placed on their computer or device. Willingness to share personal data varies between individuals — as does awareness of the potential implications of implicit agreement (or actual consent).

Businesses also have strong incentives to consider the extent to which they share the data they generate in the operation of their business with other parties. In particular, businesses have incentives to restrict access to some other parties because their data may provide them with a competitive advantage (Australian Government the Treasury 2014 p. 185). They may also be restricted in data release by legal or contractual arrangements.

Yet despite the potential market advantages, many businesses may be unaware of the usefulness of data they hold, or their data may only be useful when combined with other data that they lack the access or skills to make use of. Well under half of Australian businesses received an order over the internet in 2014 (just over half made an order in this manner) suggesting levels of digital awareness may still not be high (ABS 2015). Investigation of how effectively firms are reviewing their data holdings for opportunity to innovate (and barriers to their doing so) will be significant in this inquiry.

For some types of private sector data there are markets with willing suppliers and purchasers of data. For example, there are businesses that distribute a wide range of data from the Australian Securities Exchange and Sydney Futures Exchange trading platforms, and many businesses and individuals who purchase such data. There are also businesses that focus on collecting and selling data, such as the numerous market research companies. However, the opportunities may stretch well beyond specialist firms to a much broader cohort of firms asking themselves the question ‘what advantage does our data holding offer to us?’.

There may be some market or regulatory failures that are leading to more restrictions on the availability and use of private data than is desirable from a community-wide perspective. For example, certain types of data may have the potential to produce unanticipated spillover benefits — for example, the (US) Centers for Disease Control and Prevention uses trends on Google queries to shed light on the spread of certain infectious diseases (CDC 2009). Sharing such private sector data may generate substantial benefits to other parties, but if the private costs to the data owner of sharing the data exceed the private benefits of sharing the data, the owner may not have any incentive to share their data. However, as noted earlier, if ownership of such data is clear, mutually-beneficial arrangements for other parties to access or purchase the data may be more easily negotiated.

Further, as the Murray Inquiry (Australian Government the Treasury 2014) noted, in many cases private returns are necessary to justify investments in datasets. The challenge is ‘to maintain commercial incentives for developing datasets, while facilitating the release of data where this improves efficiency’ (p. 185).

Innovative uses of data by the private sector

As noted earlier, data can help to break down information gaps within and across parts of the economy and drive the development of new products and processes. By virtue of the enormous volume of data it generates, and its propensity for innovation and efficiency, the private sector will be the major driver of such developments.

The range of potential uses and benefits of data by the private sector is enormous and varied. Box 4 provides just one example of an innovative approach to problem solving using digital data and harnessing latent talent in the community — the Kaggle online platform.

Box 4 The Kaggle online platform and connect service

Founded in Melbourne in 2010 and subsequently relocated to San Francisco, Kaggle is a private company that hosts an online platform for data-mining and predictive-modeling competitions. Other companies can dump data with Kaggle, together with a problem description, and Kaggle’s community of computer scientists and mathematicians will then compete for prize money by posting solutions. Prize money tends to be between \$US3000 and \$US250 000, but recently a \$US3 million prize was offered by the Heritage Provider Network for predictions on hospital admissions based on historical claims data. In addition to prize money, participation in Kaggle competitions means individuals can obtain a Kaggle ranking, which has become a coveted qualification within the IT labour market.

A recent competition on cervical cancer screening provides one example of the significant benefits that can be generated through this model. A leading biotechnology company, Genentech, used Kaggle to provide a dataset of de-identified health records for data analysts. The Genentech Kaggle competition, which concluded in February 2016, resulted in the development of a more accurate predictive model for non-attendance of women for cervical cancer screenings. Identifying at-risk populations will make education and other intervention efforts more effective, ideally ultimately reducing the number of women who die from the disease.

Source: Genentech (2016).

The role of government in the availability of private sector data

There are several ways governments are involved in the availability and sharing of private data. Governments may act as the custodian of confidential private data — for example, the Australian Bureau of Statistics collects census data about individuals, confidentialises it

and releases it, and Geoscience Australia collects drilling and seismic data from resource companies and releases it after certain prescribed periods.

Governments can also have an oversight role in voluntary schemes to encourage data availability in the private sector. For example, Australia's consumer credit reporting regime is intended to promote greater data availability in the finance industry and thus enhance the decision-making capabilities of businesses in the industry, ultimately leading to gains to individual borrowers (box 5).

Box 5 Australia's consumer credit reporting regime

Consumer credit reporting involves the collection, distribution and use of credit-related data by credit reference agencies and credit providers for the purposes of assessing loan applications by potential borrowers (Australian Government 2012).

In Australia, the consumer credit reporting regime is regulated by Part IIIA of the *Privacy Act 1988*, and supported by the *Privacy Regulation 2013* and the *Privacy (Credit Reporting) Code 2014*. Information permitted to be collected and disclosed was historically limited to so-called negative events, such as defaults. However, recent reforms, which came into effect in March 2014, expanded the scope of information to include comprehensive data, such as the number and type of accounts open, the credit limit of those accounts, and credit repayment history.

Credit reporting can enhance the ability of credit providers to accurately assess which borrowers are suitable for credit, and to efficiently price credit for different individuals. It can also promote competition in credit markets by helping to level the playing field between large and small credit providers (since credit providers with large market shares already have access to detailed information about a large pool of potential borrowers) (Australian Government 2012).

However, there were also privacy concerns about credit reporting, including that credit reporting agencies will have access to large databases of personal data, and that credit data will be used for purposes other than which it is intended, such as for marketing (Australian Government 2012).

The expanded regime is voluntary, with information to be shared on a reciprocal basis (ARCA nd). The Murray Inquiry Final Report noted that the benefits of participation are likely to be relatively low for large credit providers, and that '... the pace and extent of eventual participation is not yet clear' (Australian Government the Treasury 2014).

Sources: ARCA (nd); Australian Government (2012); Australian Government the Treasury (2014).

QUESTIONS ON HIGH VALUE PRIVATE SECTOR DATA

What private sector datasets should be considered high-value data to: public policy; researchers and academics; other private sector entities; or the broader community?

In each case cited, what characteristics define such datasets?

What would be the public policy rationale for any associated government intervention?

What benefits would the community derive from increasing the availability and use of private sector data?

QUESTIONS ON ACCESS TO PRIVATE SECTOR DATA

Are there any legislative or other impediments that may be unnecessarily restricting the availability and use of private sector data? Should these impediments be reduced or removed?

What are the reasonable concerns that businesses have about increasing the availability of their data?

What principles, protocols or legislative requirements could manage the concerns of private sector data owners about increasing the availability of their data?

Should the collection, sharing and release of private sector data be standardised in some way? How could this be done and what would be the benefits and costs? What would standards that are 'fit for purpose' look like?

To what extent can voluntary data sharing arrangements — between businesses / between businesses and consumers / involving third party intermediaries — improve outcomes for the availability and use of private data? How could participation levels be increased?

Would such voluntary arrangements raise competition issues? How might this change if private sector information sharing were mandated? Is authorisation (under the Competition and Consumer Act 2010 (Cth)) relevant?

What role can governments usefully play in promoting the wider availability of private datasets that have the potential to deliver substantial spillover benefits?

How can the sharing and linking of private sector data be improved in Australia? What lessons or examples from overseas should be considered?

Who should have the ownership rights to data that is generated by individuals but collected by businesses? For which data does unclear ownership inhibit its availability and use?

Consumers' access to and control over data about them

Access to data

The Harper Review (Australian Government 2015a) noted a number of calls for consumers to have greater access to personal data that businesses have collected about them — such as individual consumers' data on their mobile phone and electricity usage, and their spending patterns at supermarkets. The Review recommended that governments, in their dealings with consumers, should 'present information and choices in ways that allow consumers to access, assess and act on them' (Australian Government 2015a, p. 305). In its response, the Australian Government supported 'allowing consumers to access information in an efficient format, especially as new technologies increase the generation of data that can improve consumer decisions but also raise consumer protection issues' (Australian Government the Treasury 2015, p. 19).

In Australia, consumers already have the right, under the *Privacy Act 1988* (Cth) (Privacy Act), to request access to their personal data held by government agencies and businesses, subject to a number of exceptions. The Privacy Act also gives individuals a right to request correction of the information held about them so that the information is accurate, up to date, complete, relevant and not misleading. Agencies must respond to a request for access within 30 days, unless impracticable, and organisations must respond within a reasonable period. If an individual is dissatisfied with the response, they may make a complaint to the entity, and then, to the Information Commissioner. Possible remedies may include taking steps to facilitate access.

However, the Privacy Act does not specify in what format the information is to be provided to consumers other than it 'must be in a manner requested by the individual if it is reasonable and practicable to do so' (Australian Government 2015a, p. 302). The Murray Inquiry report (Australian Government the Treasury 2014, p. 184) noted that a number of impediments prevent consumers from being able to access their data effectively, including little guidance on how such personal information should be provided and even confusion about what constitutes personal information.

The *Freedom of Information Act 1982* (Cth) (the FOI Act) also applies to public sector data. It gives all individuals a legally enforceable right of access to public sector documents. Access can be denied only on a ground listed in the FOI Act. The purpose of the Act was 'to open government activity to public scrutiny, so as to enhance accountability and encourage citizen engagement with public administration' (Office of the Australian Information Commissioner 2011a, p. 4).

In 2013, the Australian Government launched *myGov* as a way for individuals to more easily access government services online. The service allows individuals to access letters, statements and other types of information from a number of agencies — including Medicare, the Australian Taxation Office, Centrelink and Child Support — at a single, centralised website. Another Australian Government online data initiative is *eHealth*,

which records a range of health data on individuals that can be accessed by the relevant individual and authorised healthcare providers such as doctors and nurses — uptake, however, has been slow.

To the extent consumers value being able to access data about themselves that is held by businesses, there are likely to be some incentives — at least in competitive markets — for businesses to offer such data to their customers — and in formats that are useful to them. In principle, the less competitive the market and, arguably, the less prominent the particular business, the greater the likelihood that non-cooperation will be an issue. However, regulation may not satisfactorily address the issue — indeed, a culture of more openness may be more helpful than more regulation in generating better outcomes for consumers.

A number of voluntary, mutually-beneficial arrangements have already emerged between businesses and their customers, and these may offer useful models for wider adoption. For example, consumers are providing businesses with certain personal data in return for data-related services from the business concerned. ‘Retail tendering’, where consumers can specify the particular features they seek for various services (such as broadband and mobile services) and leave it for businesses to make offers to them, is an example of a limited data exchange with mutual benefits. There are also growing numbers of new businesses based around accessing and analysing data and selling customised services to consumers (and to other businesses).

The United Kingdom has implemented a voluntary initiative — the Midata program — that aims to provide consumers with access to data that businesses collect about their transactions in a machine-readable and re-usable format. The program’s focus is on the energy, personal current accounts, credit cards and mobile phones sectors — areas where consumers have long term, frequent interactions with service suppliers and where it may be difficult for them to compare the relative merits of product offerings from competing suppliers. In a similar vein, the United States has a ‘Smart Disclosure’ agenda to facilitate the release of public and private sector data to help consumers make more informed choices about services in energy, healthcare and finance (Australian Government 2015a).

Increasing potential role for third party intermediaries in the market place

The rapidly increasing amount of data and the pace of technology has opened business opportunities for third party intermediaries to offer tailored solutions to individuals based on their personal data. For example, an individual can provide comparison websites with certain personal data about their needs, such as their private health insurance preferences or details of the latest policy premium advice from their current insurer. A range of intermediaries also offers data aggregation and analytical services to other businesses.

However, the Murray Inquiry report noted that in many cases consumers are unable to authorise trusted third parties to access their personal information *directly from* their service provider. This reduces the ability of competitors to offer consumers better value or

tailored services, or develop advice services to better inform consumer decision making (Australian Government the Treasury 2014). ‘Data banks’ to store personal data have been suggested as an approach to allow individuals to provide access to parts of their personal information to trusted third parties (Australian Government the Treasury 2014). The Open Banking Standard in the United Kingdom is an example of such an approach (Open Data Institute 2016).

QUESTIONS ON CONSUMER ACCESS TO, AND CONTROL OVER, DATA

What impediments currently restrict consumers’ access to and use of public and private sector data about themselves? Is there scope to streamline individuals’ access to such data and, if there is, how should this be achieved?

Are regulatory solutions of value in giving consumers more access to and control over their own data?

Are there other ways to encourage greater cultural acceptance amongst businesses of consumer access to data about them?

What role do third party intermediaries currently play in assisting consumers to access and use data about themselves? What barriers impede the availability (and take-up) of services offered by third party intermediaries?

What datasets, including datasets of aggregated data on consumer outcomes at the product or provider level, would provide high value to consumers in helping them make informed decisions? What criteria should be used to identify such datasets? What, if any, barriers are impeding consumers’ access to, and use of, such data?

3 Managing the costs

While the myriad of data being generated and held by the public and private sectors provides the potential for substantial benefits, there are also legitimate interests in ensuring greater access to data does not exceed the costs of such greater access, including the potential loss of trust and privacy of individuals.

The resource costs of making data available

Increasing the availability of data is not costless. Resources are needed to ensure that data is of sufficient quality for release outside of the collecting organisation — for example, ensuring the data is well specified, consistently defined, accurate and available in a usable format. There are also costs of maintaining data over time — for example, ensuring its consistency with evolving standards and upgrading digital storage media and digital security in line with changes in technology. For many of these tasks skilled staff are required.

The public sector holds a vast array of data across multiple agencies, much of it pre-dating the digital age. This data has been collected, stored and managed in a variety of different ways (a project to create a register of all the data held by the public sector is in its early stages). Preparing this data for release in compliance with all current regulations and policies is a resource intensive task. Information management systems and data collection standards do not apply uniformly, and collection of data can be fragmented between and within agencies (Office of the Australian Information Commissioner 2012).

Government agencies lack price signals, profit motives and often any incentive to form voluntary, mutually beneficial data-sharing arrangements. The lack of market signals can also cloud decisions about how much value-adding to perform on data before releasing it, or about the format it should be in.

Where data is made available, it is often provided on a one-off and ‘destroy after use’ basis. The considerable effort undertaken in cleaning and documenting such data is often lost. This is particularly the case for data made available for research purposes. For instance, even when a researcher’s discoveries about the data quality and quirks are relayed back to the data custodian, the often considerable effort to improve the functionality of the dataset goes to waste once the dataset is deleted.

As noted by the Harper Review (Australian Government 2015a), just providing information on the performance of providers or products is not enough to guarantee that good choices will be made by consumers. Similarly, making available vast volumes of public sector data is not necessarily in the best interests of the community. The objective of increasing data access should not simply be to increase the volume of available data. Rather, it should be to increase access to that data for which wider availability would be expected to deliver benefits to the community greater than the costs of making it available — including the potential costs of reducing individuals’ privacy, security breaches and resource costs.

QUESTIONS ON RESOURCE COSTS OF ACCESS

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

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

What pricing principles should be applied to different datasets? What role should price signals play in the provision of public sector data?

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

Trust

For the economic benefits of data to be fully realised, it will be essential to maintain individuals' and businesses' confidence and trust in how data is collected, stored and used. Trust plays a central role in social and economic interactions, enabling individuals and businesses to engage with confidence that the outcomes of their interactions will be mutually beneficial. It may also increase the willingness of individuals to provide information about themselves that is accurate and true, thus increasing the value of that information.

Some reports (for example, World Economic Forum and Bain and Company 2011) have noted that many individuals do not understand how their personal information is currently being collected and used and, should they find out, may lose trust and stop using services that collect their personal information.

Individuals and businesses are required to provide governments with significant amounts of data in order to access government services or meet regulatory requirements (for example, to comply with taxation requirements). Similarly, providing data to private institutions is sometimes a necessary precondition for participation in markets (for example, opening a bank account generates financial transaction data and using a telephone generates call records). The Public Sector Data Management framework recognises that:

... it is crucial that we [the public sector] have the trust of the public. Strong assurances about data privacy and security based on rigorous adherence to protocols, and demonstrated value, are key. (Australian Government Department of the Prime Minister and Cabinet 2015, p. 32)

Privacy

A considerable proportion of the data being collected, stored and transmitted, increasingly electronically, consists of personal data about individuals, some of it potentially sensitive and which the individuals concerned may, legitimately, not wish to have distributed widely. Globally there is growing debate over how societies should consider privacy against the economic benefits associated with the rapidly growing volume of data being generated and used.

The economic consequences of information sharing can be either welfare enhancing or reducing:

In choosing the balance between sharing or hiding one's personal information (and in choosing the balance between exploiting or protecting individuals' data), individuals and organizations face complex, sometimes intangible, and often ambiguous trade-offs. Individuals want to protect the security of their data and avoid the misuse of information they pass to other entities. However, they also benefit from sharing with peers and third parties information that makes mutually satisfactory interactions possible. Organizations want to know more about the parties they interact with, tracking them across transactions. Yet, they do not want to alienate those parties with policies that may be deemed too invasive. (Acquisti 2010, p. 3)

There are challenges in managing individual privacy. For instance, much of the privacy-related regulation is based around the notion of ‘consent’. However, consent to the collection of data may not always be genuine, informed or meaningful (Acquisti 2010). Further, as noted earlier, while some of this data is ‘volunteered’ by individuals — that is, knowingly and willingly provided — an increasing proportion of data is ‘observed’. Observed data includes the online tracking of individuals and the collection and analysis of related personal information — and data that is collected incidentally as a byproduct of other activities — for example, images of people, with date and time, collected by CCTV systems intended to record licence plates and establish who should pay road tolls.

It is also difficult for individuals to know exactly what information is being collected about them, and what it is being used for. Advances in data analytics are making it increasingly easy to generate inferences about individuals using data collected in different contexts. With sufficient data, analysts can predict, with varying degrees of accuracy, the likelihood that an individual will possess certain characteristics or undertake certain actions. Concerns have been raised that the information inferred through data analytics could be used to exploit the vulnerabilities and receptiveness of individuals (OECD 2015a). Linked, de-identified datasets can also be used to re-identify individuals, and this risk tends to increase as more datasets are available for analysis.

Digitisation of media, increasing use of social media platforms and the dominance of search aggregators such as Google also enable personal information about an individual to be shared more easily with a wider audience and be more readily searchable. This can be particularly problematic where this information is deeply private or prejudicial to an individual. One issue is whether individuals should be able to request deletion of information about themselves. There can be strong public interest, freedom of expression, and other legal and practical grounds for allowing particular organisations to retain such information, and it can also be costly and complicated to delete, particularly when third parties are involved.

Existing protections to privacy

Governments have a role to play in upholding societal norms on privacy protection. Legal protection of privacy in Australia currently comprises a mix of Commonwealth, state and territory legislation that regulates the collection, use and disclosure of individuals’ personal information.

The Privacy Act and the associated Australian Privacy Principles apply to private sector entities with an annual turnover of at least \$3 million and to all Commonwealth Government agencies, subject to certain exemptions.

In 2008, the Australian Law Reform Commission (ALRC) concluded a comprehensive review of Australia’s privacy laws. The ALRC investigated the extent to which the Privacy Act and related laws continue to provide an effective framework for the protection of privacy in Australia, having regard to the need of individuals for privacy protection in an

evolving technological environment, and the desirability of minimising the regulatory burden on businesses in this area.

Reforms stemming from the review included the creation of the Australian Privacy Principles and provision of greater clarity on how individuals can access and control their personal data. A subsequent review by the ALRC in 2014 considered prevention of and remedies for serious invasions of privacy in the digital era. There have also been a number of privacy reviews by state law reform commissions in recent years. Australian states and territories (except for Western Australia and South Australia) have their own information privacy legislation that applies to their agencies (and the interactions of private businesses with these agencies):

- *Privacy and Personal Information Protection Act 1998* (NSW)
- *Privacy and Data Protection Act 2014* (Vic)
- *Information Privacy Act 2009* (Qld)
- *Personal Information Protection Act 2004* (Tas)
- *Information Privacy Act 2014* (ACT)
- *Information Act 2002* (NT).

Further, separate Commonwealth, state and territory legislation governs the collection, use and disclosure of some specific types of information. In New South Wales, for example, the *National Health Act 1953* (Cth) and the *Health Records and Information Privacy Act 2002* (NSW) govern the collection and use of health data. In addition, substantial amounts of data on individuals may be collected through government programs — for example, data collected for the purposes of monitoring program outcomes. There may also be a range of agency-specific legislation, policies and procedures (for example, the requirement to obtain ethics committee approval) that govern the collection, use and disclosure of this information.

Rapid technological change adds an additional layer of complexity. As recognised by the ALRC, it is important that privacy laws be:

... sufficiently flexible to adapt to rapidly changing technologies and capabilities without the need for constant amendments. At the same time, they should be drafted with sufficient precision and definition to promote certainty as to their application and interpretation. (Australian Law Reform Commission 2014, p. 32)

The Australian Privacy Principles do not impose technology-specific obligations. Additionally, the Privacy Act requires the Office of the Australian Information Commissioner to undertake research into, and monitor developments in, data processing and technology (including data matching and linkage) to ensure that any adverse effects of such developments on the privacy of individuals are minimised, and report to the Minister the results of that research and monitoring.

Individuals do not currently have the right under the Australian Privacy Principles to delete information about themselves. However, the Australian Privacy Principles require entities holding personal information about an identifiable individual to destroy or de-identify that information when it is no longer required for a specific purpose and an Australian Senate Committee recently proposed laws criminalising the non-consensual sharing of intimate information about an individual (Legal and Constitutional Affairs References Committee, Australian Senate 2015).

Despite the recent reforms to the Privacy Act, having multiple pieces of legislation governing privacy can make it difficult and time-consuming for agencies and businesses to understand and fulfil their obligations. The complex governance of personal information can also make it difficult for individuals to understand and act on their rights.

QUESTIONS ON PRIVACY PROTECTION

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

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

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

Are further changes to the privacy-related policy framework needed? What are these specific changes and how would they improve outcomes? Have such approaches been tried in other jurisdictions?

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

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

What lessons from overseas jurisdictions can Australia learn from regarding the use of individuals' and businesses' data, particularly in regard to protecting privacy and commercially sensitive or commercially valuable information?

What are the benefits and costs of allowing an individual to request deletion of personal information about themselves? In what circumstances and for what types of information should this apply?

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

Other restrictions around the release of particular data

Some agencies and organisations have restrictions on the release and use of particular data for a range of reasons such as commercial-in-confidence, stability of institutional

structures, or national security. These restrictions are typically embedded in agency-specific regulations, policies and protocols, such as the *Offshore Petroleum and Greenhouse Gas Storage (Resource Management and Administration) Regulations 2011* (Cth). These variations in treatment can add to the complexity surrounding data access and use. Some data may be collected as part of an agency's regulatory function — for instance, under the *Anti-Money Laundering and Counter-Terrorism Financing Act 2006* (Cth), AUSTRAC collects transaction data from entities such as banks bound to report under the legislation.

There is also legislation governing the collection, use and disclosure of data for national security and law enforcement purposes, such as the *Intelligence Services Act 2001* (Cth) and the *Australian Security Intelligence Organisation Act 1979* (Cth).

QUESTIONS ON OTHER RESTRICTIONS

Having regard to current legislation and practice, are further protocols or other measures required to facilitate the disclosure and use of data about individuals while protecting privacy interests? What form should any such protocols or other measures take?

Is there need for a more uniform treatment of commercial-in-confidence data held by the Australian Government and state and territory governments?

Are there merits in codifying the treatment and classification of business data for privacy or security purposes? What would this mean in practice?

Data security

Being a valuable resource, data is often a target for theft or misuse by organised crime groups and individual hackers. Data is also sometimes released unintentionally — for example, through human error or technical malfunctions. Where personal data is collected, stored or processed — by businesses or government agencies — security incidents can adversely affect individuals' privacy and potentially impose significant costs on the individuals concerned.

Data breaches are incidents in which sensitive, protected or confidential data is viewed, stolen or used by an individual unauthorized to do so. Aside from the threats they present to the privacy of individuals, breaches can compromise the operations of the business or agency concerned (in the case of commercially valuable or protected information) and potentially tarnish the reputation of the business or government agency that is breached.

Agencies and organisations have obligations under the Privacy Act to put in place security safeguards and take reasonable steps to protect the personal data that they hold from misuse, interference and loss, and from unauthorised access, modification and disclosure. In addition to the requirements under the Privacy Act to protect personal data, there is a range of policies relevant to public sector information security. For example, the Australian

Signals Directorate has issued the 2015 Australian Government Information Security Manual, which sets standards for the security of government ICT systems.

The Office of the Australian Information Commissioner administers a voluntary data breach notification scheme, and received 110 notifications in 2014-15, a 64 per cent increase on the number received in 2013-14 (Office of the Australian Information Commissioner 2015a). The draft Privacy Amendment (Notification of Serious Data Breaches) Bill proposes to impose a legal obligation to notify an individual if their data has been compromised. The scheme could apply to any agency or business that is currently subject to the Privacy Act.

QUESTIONS ON DATA SECURITY

Are security measures for public sector data too prescriptive? Do they need to be more flexible to adapt to changing circumstances and technologies?

How do data security measures interact with the Privacy Act?

How should the risks and consequences of public sector and private sector data breaches be assessed and managed? Is data breach notification an appropriate and sufficient response?

References

- ABS (Australian Bureau of Statistics) 2015, *Summary of IT Use and Innovation in Australian Business, 2013-14*, Cat. no. 8166.0, Canberra.
- Acquisti, A. 2010, *The Economics of Personal Data and the Economics of Privacy*. Background Paper #3, OECD Roundtable on The Economics of Personal Data and Privacy: 30 Years after the OECD Privacy Guidelines, December.
- ARCA (Australian Retail Credit Association) nd, Principles of Reciprocity and Data Exchange, <http://www.arca.asn.au/focus/principles-of-reciprocity-data-exchange-prde.html> (accessed 4 February 2016).
- Australian Government 2012, Explanatory Memorandum — Privacy Amendment (Enhancing Privacy Protection) Bill 2012, Canberra.
- Australian Government the Treasury 2014, *Financial System Inquiry Final Report* (Murray Inquiry), Canberra.
- 2015, *Australian Government Response to the Competition Policy Review*, Canberra.
- Australian Government 2014, *Towards Responsible Government — The Report of the National Commission of Audit, Phase One*, Canberra.
- 2015a, *Competition Policy Review, Final Report* (Harper Review) Canberra.
- 2015b, *Public Data Policy Statement*, Canberra.
- Australian Government Department of the Prime Minister and Cabinet 2015, *Public Sector Data Management*, Canberra.
- Australian Government Department of Communications and the Arts Bureau of Communications Research 2016, *Open government data and why it matters — A critical review of studies on the economic impact of open government data*, Canberra.
- Australian Law Reform Commission 2014, *Serious invasions of privacy in the digital era*, ALRC Report 123, Sydney.
- Brynjolfsson, E. and McAfee, A. 2014, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W. W. Norton & Company.
- CDC (Centers for Disease Control and Prevention) 2009, *More Diseases Tracked by Using Google Trends* http://wwwnc.cdc.gov/eid/article/15/8/09-0299_article (accessed 7 April 2016).
- EMC Corporation 2014, *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things*, Executive Summary, <http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm> (accessed 22 February 2015).

-
- Garvan Institute of Medical Research 2015, *Garvan Partners with Genomics England*, <http://www.garvan.org.au/news/news/garvan-partners-with-genomics-england> (accessed 26 February 2016).
- Genentech 2016, *Help prevent cervical cancer by identifying at-risk populations*, <https://www.kaggle.com/c/cervical-cancer-screening> (accessed 13 April 2016).
- Houghton, J. 2011, *Costs and Benefits of Data Provision*, Report to the Australian National Data Service, Centre for Strategic Economic Studies Victoria University.
- IBM 2016, *What is big data?* <http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html> (accessed 26 February 2016).
- Legal and Constitutional Affairs References Committee, Australian Senate 2015, http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Legal_and_Constitutional_Affairs/Revenge (accessed 8 April 2016).
- McKinsey Global Institute 2013, *Open Data: Unlocking innovation and performance with liquid information*.
- New York Times 2013, *Is Big Data an Economic Big Dud?*, http://www.nytimes.com/2013/08/18/sunday-review/is-big-data-an-economic-big-dud.html?pagewanted=1&_r=2&emc=eta1 (accessed 6 April 2016).
- OECD (Organisation for Economic Cooperation and Development) 2015a, *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris.
- 2015b, *Digital Government, Government at a Glance 2015*, OECD Publishing, Paris.
- 2015c, *Health Data Governance: Privacy, Monitoring and Research*, Policy Brief, October, Paris.
- Office of the Australian Information Commissioner 2011a, *Guide to the Freedom of Information Act 1982*, Canberra.
- 2011b, *Report on review and development of principles*, Canberra.
- 2012, *Open public sector information: from principles to practice. Report on agency implementation of the Principles on open public sector information*, February.
- 2015a, *Annual Report 2014-15*, Canberra.
- 2015b, *Guide to securing personal information*, <https://www.oaic.gov.au/agencies-and-organisations/guides/guide-to-securing-personal-information> (accessed 15 March 2016).
- Open Data Institute 2016, *The Open Banking Standard*, <http://theodi.org/open-banking-standard> (accessed 24 March 2016).
- Open Knowledge 2015, *Global Open Data Index*, <http://index.okfn.org/place/> (accessed 30 March 2016).
- PricewaterhouseCoopers 2014, *Deciding with data — How data-driven innovation is fuelling Australia's economic growth*.

-
- PC (Productivity Commission) 2010a, *Annual Report 2009-10*, Annual Report Series, Productivity Commission, Canberra.
- 2010b, *Contribution of the Not-for-Profit Sector*, Research Report, Canberra.
- 2010c, *Gambling*, Report no. 50, Canberra.
- 2011a, *Caring for Older Australians*, Report No. 53, Final Inquiry Report, Canberra.
- 2011b, *Disability Care and Support*, Report no. 54, Canberra.
- 2012a, *Barriers to Effective Climate Change Adaptation*, Report No. 59, Final Inquiry Report, Canberra.
- 2012b, *Impacts of COAG Reforms: Business Regulation and VET*, Research Report, Volume 1 – Overview, Canberra.
- 2012c, *Regulatory Impact Analysis: Benchmarking*, Research Report, Canberra.
- 2012d, *Schools Workforce*, Research Report, Canberra.
- 2013a, *Annual Report 2012-13*, Annual Report Series, Productivity Commission, Canberra.
- 2013b, *Regulator Engagement with Small Business*, Research Report, Canberra.
- 2014a, *Childcare and Early Childhood Learning*, Inquiry Report No. 73, Canberra.
- 2014b, *Public Infrastructure*, Inquiry Report No. 71, Canberra.
- 2015a, *Efficiency in Health*, Commission Research Paper, Canberra.
- 2015b, *Housing Assistance and Employment in Australia*, Commission Research Paper, Canberra.
- PSMA Australia 2016, *PSMA responds to Australian Government's release of G-NAF to the public*, <https://www.pasma.com.au/publications/news/psma-responds-australian-governments-release-g-naf-public> (accessed 12 April 2016).
- Tanner, L. 2010, *Declaration of Open Government*, <http://www.finance.gov.au/blog/2010/07/16/declaration-open-government/> (accessed 21 March 2016).
- World Economic Forum and Bain and Company 2011, *Personal Data: The Emergence of a New Asset Class*, World Economic Forum.
- World Wide Web Foundation 2015, *Open Data Barometer Global Report*, 2nd edition.

Attachment A — How to make a submission

This Commission invites interested people and organisations to make a written submission.

How to prepare a submission

Submissions may range from a short letter outlining your views on a particular topic to a much more substantial document covering a range of issues. Where possible, you should provide evidence, such as relevant data and documentation, to support your views.

Generally

- Each submission, except for any attachment supplied in confidence, will be published on the Commission's website shortly after receipt, and will remain there indefinitely as a public document.
- The Commission reserves the right to not publish material on its website that is offensive, potentially defamatory, or clearly out of scope for the inquiry or study in question.

Copyright

- Copyright in submissions sent to the Commission resides with the author(s), not with the Commission.
- Do not send us material for which you are not the copyright owner — such as newspaper articles — you should just reference or link to this material in your submission.

In confidence material

- This is a public review and all submissions should be provided as public documents that can be placed on the Commission's website for others to read and comment on. However, information which is of a confidential nature or which is submitted in confidence can be treated as such by the Commission, provided the cause for such treatment is shown.
- The Commission may also request a non-confidential summary of the confidential material it is given, or the reasons why a summary cannot be provided.
- Material supplied in confidence should be clearly marked 'IN CONFIDENCE' and be in a separate attachment to non-confidential material.
- You are encouraged to contact the Commission for further information and advice before submitting such material.

Privacy

- For privacy reasons, all personal details (e.g. home and email address, signatures, phone, mobile and fax numbers) will be removed before they are published on the website. Please do not provide a these details unless necessary.
- You may wish to remain anonymous or use a pseudonym. Please note that, if you choose to remain anonymous or use a pseudonym, the Commission may place less weight on your submission.

Technical tips

- The Commission prefers to receive submissions as a Microsoft Word (.docx) files. PDF files are acceptable if produced from a Word document or similar text based software. You may wish to research the Internet on how to make your documents more accessible or for the more technical, follow advice from Web Content Accessibility Guidelines (WCAG) 2.0<<http://www.w3.org/TR/WCAG20/>>.
- Do not send password protected files.
- Track changes, editing marks, hidden text and internal links should be removed from submissions.
- To minimise linking problems, type the full web address (for example, <http://www.referred website.com/folder/file name.html>).

How to lodge a submission

Submissions should be lodged using the online form on the Commission's website. Submissions lodged by post should be accompanied by a submission cover sheet.

Online* **www.pc.gov.au/inquiries/current/data-access**

Post* Data Availability and Use
 Productivity Commission
 GPO Box 1428
 Canberra City ACT 2601

* If you do not receive notification of receipt of your submission to the Commission, please contact the Administrative Officer.

Due date for submissions

Please send submissions to the Commission by **Friday 29 July 2016**.