



Australian Government
Productivity Commission

Data Availability and Use

Productivity Commission
Inquiry Report
Overview & Recommendations

No. 82, 31 March 2017

© Commonwealth of Australia 2017

ISSN 1447-1337 (online)

ISSN 1447-1329 (print)

ISBN 978-1-74037-619-8 (overview online)

ISBN 978-1-74037-618-1 (overview print)



Except for the Commonwealth Coat of Arms and content supplied by third parties, this copyright work is licensed under a Creative Commons Attribution 3.0 Australia licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/3.0/au>. In essence, you are free to copy, communicate and adapt the work, as long as you attribute the work to the Productivity Commission (but not in any way that suggests the Commission endorses you or your use) and abide by the other licence terms.

Use of the Commonwealth Coat of Arms

For terms of use of the Coat of Arms visit the 'It's an Honour' website: <http://www.itsanhonour.gov.au>

Third party copyright

Wherever a third party holds copyright in this material, the copyright remains with that party. Their permission may be required to use the material, please contact them directly.

Attribution

This work should be attributed as follows, *Source: Productivity Commission, Data Availability and Use, Inquiry Report.*

If you have adapted, modified or transformed this work in anyway, please use the following, *Source: based on Productivity Commission data, Data Availability and Use, Inquiry Report.*

An appropriate reference for this publication is:

Productivity Commission 2017, *Data Availability and Use: Overview & Recommendations*, Report No. 82, Canberra

Publications enquiries

Media and Publications, phone: (03) 9653 2244 or email: maps@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).



Australian Government
Productivity Commission

Canberra Office

Level 2, 4 National Circuit
Barton ACT 2600

GPO Box 1428
Canberra City ACT

Telephone 02 6240 3200

Melbourne Office

Telephone 03 9653 2100

www.pc.gov.au

31 March 2017

The Hon Scott Morrison MP
Treasurer
Parliament House
CANBERRA ACT 2600

Dear Treasurer

In accordance with Section 11 of the *Productivity Commission Act 1998*, we have pleasure in submitting to you the Commission's final report into *Data Availability and Use*.

Yours sincerely

A handwritten signature in black ink, appearing to read 'Peter Harris'.

Peter Harris
Chair

A handwritten signature in black ink, appearing to read 'Melinda Cilento'.

Melinda Cilento
Commissioner

Terms of reference

I, Scott Morrison, Treasurer, pursuant to Parts 2 and 3 of the *Productivity Commission Act 1998*, hereby request that the Productivity 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
 - (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 and 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
 - (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
- (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
 - (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.

Scott Morrison
Treasurer

[Received 21 March 2016]

Disclosure of interests

The *Productivity Commission Act 1998* specifies that where Commissioners have or acquire interests, pecuniary or otherwise, that could conflict with the proper performance of their functions during an inquiry they must disclose the interests.

Ms Cilento has advised the Commission that she is a director of Australian Unity (which made a submission to this Inquiry) and of Woodside Petroleum (which is referred to in an included example in the Report). Ms Cilento is also co-chair of the National Australia Bank's Advisory Council on Corporate Responsibility.

Contents

Terms of reference	v
Acknowledgments	x
Abbreviations	xi
Overview	1
Key points	2
Access denied – Australia’s lost opportunities	4
Fundamental change is needed	12
The new Framework	14
Implementation	29
Findings and recommendations	33
Key points from chapters	53
Chapter 1 Australia’s data landscape	53
Chapter 2 Opportunities enabled by data	54
Chapter 3 What holds us back?	55
Chapter 4 A way forward: what we must aim at	56
Chapter 5 New competition policy — a right to use your data	57
Chapter 6 Sharing and releasing data for community benefits	58
Chapter 7 Getting value from Australia's national interest datasets	59
Chapter 8 A modernised regulatory framework	60
Chapter 9 Transformation and pricing decisions	61
Chapter 10 Implementing the new data Framework	62
References	63

Acknowledgments

The Commissioners express their appreciation to the staff who worked on the Inquiry report and underlying analysis.

The Inquiry team was led by Rosalyn Bell and included Elise Whalan, Miriam Veisman-Apter, Greg Thompson, Joshua Runciman, Claire Prideaux, and Anthony Housego. Ian Moran, Kathryn Ovington (from Attorney-General's Department) and Gavin Walker (from CSIRO Data61) were also a part of the team for completion of the draft Report.

Abbreviations

ABS	Australian Bureau of Statistics
ACCC	Australian Competition and Consumer Commission
ACMA	Australian Communications and Media Authority
AGD	Attorney-General's Department
AGIMO	Australian Government Information Management Office
AGLDWG	Australian Government Linked Data Working Group
AIHW	Australian Institute of Health and Welfare
ALRC	Australian Law Reform Commission
ANAO	Australian National Audit Office
ANDS	Australian National Data Service
API	Application Programming Interface
APP	Australian Privacy Principle
APRA	Australian Prudential Regulation Authority
ARA	Accredited Release Authority
ARC	Australian Research Council
ARCA	Australian Retail Credit Association
ASAC	Australian Statistics Advisory Council
ASIC	Australian Securities and Investments Commission
ATO	Australian Tax Office
AURIN	Australian Urban Research Infrastructure Network
AUSTRAC	Australian Transactions Reporting and Analysis Centre
BLADE	Business Longitudinal Analytical Data Environment
CCR	Comprehensive Credit Reporting
COAG	Council of Australian Governments
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DFAT	Department of Foreign Affairs and Trade
DHS	Department of Human Services
DPMC	Department of Prime Minister and Cabinet
DSS	Department of Social Services
DTA	Digital Transformation Agency

DTO	Digital Transformation Office
FOI	Freedom of Information
GIF	Graphics Interchange Format
GIS	Geographic Information System
G-NAF	Geocoded National Address File
GPS	Global Positioning System
HILDA	Household, Income and Labour Dynamics Australia
HREC	Human Research Ethics Committee
ICT	Information and Communications Technology
IDI	Integrated Data Infrastructure
IP	Internet Protocol
IT	Information Technology
IoT	Internet of Things
JSON	JavaScript Object Notation
MADIP	Multi-Agency Data Integration Project
MBS	Medicare Benefits Schedule
MOG	Machinery of Government
MOU	Memorandum of Understanding
NAA	National Archives Australia
NCRIS	National Collaborative Research Infrastructure Strategy
NDC	National Data Custodian
NHMRC	National Health and Medical Research Council
NID	National Interest Dataset
NSS	National Statistical Service
NSW DAC	New South Wales Data Analytics Centre
OAIC	Office of the Australian Information Commissioner
OECD	Organisation for Economic Co-operation and Development
PBS	Pharmaceutical Benefits Scheme
PC	Productivity Commission
PHRN	Population Health Research Network
RBA	Reserve Bank of Australia
SURE	Secure Unified Research Environment
WWWF	World Wide Web Foundation

OVERVIEW

Key points

- Extraordinary growth in data generation and usability has enabled a kaleidoscope of new business models, products and insights. Data frameworks and protections developed prior to sweeping digitisation need reform. This is a global phenomenon and Australia, to its detriment, is not yet participating.
- Improved data access and use can enable new products and services that transform everyday life, drive efficiency and safety, create productivity gains and allow better decision making.
- The substantive argument for making data more available is that opportunities to use it are largely unknown until the data sources themselves are better known, and until data users have been able to undertake discovery of data.
- Lack of trust by both data custodians and users in existing data access processes and protections and numerous hurdles to sharing and releasing data are choking the use and value of Australia's data. In fact, improving trust community-wide is a key objective.
- Marginal changes to existing structures and legislation will not suffice. Recommended reforms are aimed at moving from a system based on risk aversion and avoidance, to one based on transparency and confidence in data processes, treating data as an asset and not a threat. Significant change is needed for Australia's open government agenda and the rights of consumers to data to catch up with achievements in competing economies.
 - At the centre of recommended reforms is a new Data Sharing and Release Act, and a National Data Custodian to guide and monitor new access and use arrangements, including proactively managing risks and broader ethical considerations around data use.
 - A new Comprehensive Right for consumers would give individuals and small/medium businesses opportunities for active use of their own data and represent fundamental reform to Australia's competition policy in a digital world. This right would create for consumers:
 - powers comparable to those in the Privacy Act to view, request edits or corrections, and be advised of the trade to third parties of consumer information held on them
 - a new right to have a machine-readable copy of their consumer data provided either to them or directly to a nominated third party, such as a new service provider.
- A key facet of the recommended reforms is the creation of a data sharing and release structure that indicates to all data custodians a strong and clear cultural shift towards better data use that can be dialled up for the sharing or release of higher-risk datasets.
 - For datasets designated as national interest, all restrictions to access and use contained in a variety of national and state legislation, and other program-specific policies, would be replaced by new arrangements under the Data Sharing and Release Act. National Interest Datasets would be resourced by the Commonwealth as national assets.
 - A suite of Accredited Release Authorities would be sectoral hubs of expertise and enable the ongoing maintenance of, and streamlined access to, National Interest Datasets as well as to other datasets to be linked and shared or released.
 - A streamlining of ethics committee approval processes would provide more timely access to identifiable data for research and policy development purposes.
- Incremental costs of more open data access and use — including those associated with better risk management and alterations to business data systems — will exist but should be substantially outweighed by the opportunities presented.
- Governments that ignore potential gains through consumer data rights will make the task of garnering social licence needed for other data reforms more difficult. Decoupling elements of this Framework runs the risk of limiting benefits to, and support from, the wider public.

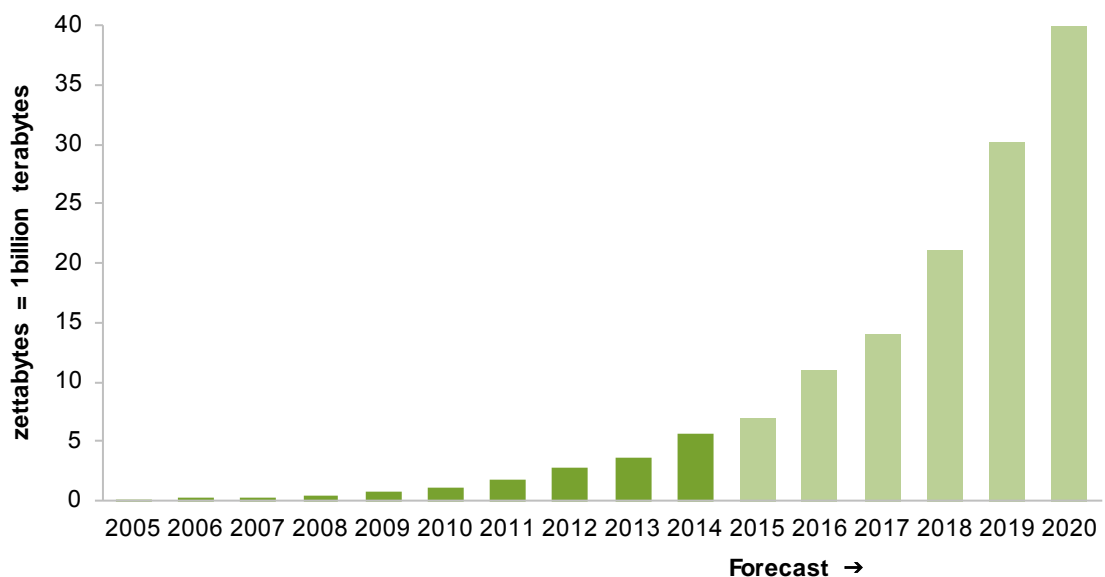
Overview

Thirty years ago, data for most people was primarily about details on paper. Data was largely collected and stored on paper (encyclopaedias, forms, bills, bank deposit slips and phone books); mail actually meant a letter in the letter box. Access to data was clear and locational (you needed keys to the filing cabinet); as was its destruction (via a shredding machine). With the mass digitisation of data, the capacity to collect data through everyday Internet activity and transactions, and through technologies such as sensors, cameras and mobile devices, means that what is ‘data’, and who can or should have a say in how it is collected, stored, transformed and used is no longer so simple.

Until this Inquiry, there has been no structured attempt to comprehensively review this matter in Australia, despite the magnitude of the transformation under way.

Data now includes material (raw or processed) on: the characteristics, status, appearance or performance of an individual, product or service, or object (including infrastructure and environmental assets); and expressed or inferred opinions and preferences. The generation of data is seemingly heading upward on an unbounded trajectory (figure 1).

Figure 1 **Data generated (global)**



Source: United Nations Economic Commission for Europe (2015)

By some estimates, the amount of digital data generated globally in 2002 (five terabytes) is now generated every two days, with 90% of the world's information generated in just the past two years (IBM 2016). As we are now only in the very nascent stage of the Internet of Things (whereby our business equipment, vehicles, appliances and wearable devices can communicate with each other and generate data), the upward trend in data generated is more likely than not to accelerate into the future.

Falling costs (per record) of digital data storage, and the spread of low-cost and powerful analytics tools and techniques to extract patterns, correlations and interactions from within data, are also making data analytics more usable and valuable. Yet much of the data being generated remains underutilised. Some estimate that up to 80% of data generated globally may prove to have no value (numerous duplicative digital photos, for example). But still, less than 5% of the potentially useful data is actually analysed to generate information, build knowledge, and thus inform decision making and action (EMC Corporation 2014). And some data that was previously of limited value is becoming valuable as new uses for it emerge, analytical capabilities improve, new linkages are established, or investments made to improve its quality. There is enormous untapped potential in Australia's data.

Access denied — Australia's lost opportunities

With technological developments and advances in analytical techniques, not only is the volume of data being generated and collected growing, but so too is the scope to make use of data in innovative ways in every sphere of life.

Increased access to data can facilitate the development of ground-breaking new products and services that fundamentally transform everyday life. Many are widely known — apps that tell you in real time where to find vacant car parking places, the fastest route to travel to the city at the time you want to go, or which electricity provider may offer you a deal based on your pattern of energy use, are all examples that rely on data analysis.

But better access to and use of data can also benefit business and government through improved operational processes and productivity. Examples abound of new found opportunities — in supply chain logistics, saving time and money; through more cost effective infrastructure and machinery maintenance and planning; improved safety and efficiency in aircraft engines; and in the capacity to better respond to and manage emergencies. And data is critical to building the evidence base to underpin incremental improvements, allowing governments and businesses to offer products and services that are more customised, coordinated or timely. The potential value of data is tremendous; as is the scope for Australia to forgo much of this value under the misconception that denial of access minimises risks.

While this Report highlights some examples of where data is already being used to benefit the community, these are the tip of the iceberg. What is already being done with data overseas is indicative of what is possible in Australia, if only more data could be released for use and the risks managed.

Health data exemplifies the problem

Australia's health sector exemplifies many of these opportunities, to date largely foregone, due to impediments and distrust around data use (box 1). Data from the sector that could be more widely used includes:

- broad level performance data on expenditure and activity at particular medical facilities (the number of available public and private hospital beds by State and Territory) or for particular medical conditions (the number of people diagnosed with asthma in each of the past 20 years and public expenditure on particular types of asthma treatment)
- finer level performance data on particular parts of the sector (the number of serious complications following orthopaedic surgery at each hospital, or how drugs prescribed for particular medical conditions vary across medical practitioners);
- data from the health records of individual patients (documented reasons for visits to health professionals, the results from diagnostic testing undertaken, prescriptions received, private and public health insurance claimed)
- data collected through personally controlled devices, such as smartphones and health monitors, that have increasing potential to assist both medical practitioners and patients.

From the Commission's experience with its annual *Report on Government Services*, data that allows performance monitoring and comparison of government activities is a fundamental starting point for improving the delivery of those activities to the community. While data in that publication motivates a closer examination of practices within particular sectors and jurisdictions, the highly aggregated level limits its use by governments, businesses and the community in making better informed decisions about health products and services. Yet behind many of these thousands of aggregated data points are powerful datasets, the equivalent of which capable, trusted researchers in other nations — the United States, New Zealand and the United Kingdom — can and do actively analyse to enable discovery and solutions to seemingly intractable problems. In that context, we fall short. Some of our best health researchers use UK health datasets, as ours are unavailable to them. Others wait up to eight years for access, in areas of life-saving significance.

Inquiry participants highlighted a range of health sector data that could underpin substantial long lasting benefits for the Australian community. We cannot afford to ignore these.

Using data to anticipate and prepare for community and individual health needs

Health data can help policy makers and researchers to:

- identify emerging health issues within communities and factors that contribute to particular medical conditions;
- assess the safety of pharmaceuticals and other treatment options on an ongoing basis; and
- evaluate the effectiveness and efficiency of health policy.

Box 1**Australia's health data — an underutilised resource that could be saving lives**

Due to a multitude of legal, institutional and technical reasons, Australia stands out among other developed countries as one where health information is poorly used (OECD 2015b):

The health sector is very good at generating and storing data. It is less effective at translating this data into useful information. It is poor at linking and sharing information between health professionals, where it could be used to improve health outcomes and system efficiency. Worst of all is the health sector's ability and willingness to share data with consumers (Medibank Private, sub. 98. p. 2).

The implications of this situation are significant. At the individual level, patients are required in many cases to act as information conduits between the various health care providers they see. Inadequate information can lead to errors in treating patients (Joint Council of Social Service Network, sub. 170). At the system level, inefficient collection and sharing leads to data gaps and unnecessary expenditure:

In a clinical sense, the lack of effective data sharing or data integration between different healthcare providers means that, in many cases, they are not in a position to deliver fully informed healthcare at an episodic level nor provide effective continuity-of-care to consumers. ... [T]his represents a considerable waste of time and effort resulting in ongoing data quality issues. At a deeper level, it means that providers do not have access to a fully integrated patient history, with the potential for unnecessary duplication of services, particularly pathology and radiology services. This inability to see the 'big picture' of a patient's health is potentially dangerous, especially in an emergency setting....

Administrative [health] data collections contain incomplete filtered and filleted data that limit their usefulness for planning and managing precision population health and monitoring the efficiency and quality of health services (Srinivasan et al. 2016, pp. 12, 21).

Furthermore, the lengthy approval process for researchers requesting access to personal data limits their ability to make potentially life-saving discoveries:

- Nearly five years after requesting the data, researchers at the University of Melbourne received de-identified information about CT scans and cancer notifications. Their work showed there was an increased cancer risk for young people undergoing CT scans, and led to changes in medical guidelines for the use of scans. "Had [the] study been approved sooner, and been able to proceed at an earlier date..., we would have had results sooner, with potential benefits in terms of improved guidelines for CT usage, lesser exposures and fewer cancers" (John D Mathews, sub. 36, p. 13).
- Since 2008, government agencies have been providing funding to the Vaccine Assessment Using Linked Data Safety Study. Among other objectives, this study examines whether there is a relationship between vaccination and admission to hospital or death. The study requires data from both the Australian and State Governments. Obtaining data from the Australian Government has taken six and a half years; State data has not yet been linked (Research Australia, sub. 117). The researchers have been waiting for the linked data for more than eight years.

In the United Kingdom, administrative hospital records linked (via unique patient health service number) with a number of cancer screening registries have been used to improve how and when cancer is diagnosed (to increase early detection and survival). Undertaking similar analysis in Australia would require linking of data held by a range of groups, including data from Medicare Australia, the Australian Government Department of Health

and its counterparts in the States and Territories, various cancer registries and other organisations.

There is already strong support for using Australia's health data in research. A recent survey revealed that over 90% of Australians were willing to share their de-identified health data to advance medical research and improve patient care (Research Australia 2016). Yet more effective use of data is not being sufficiently enabled. Inquiry participants noted a wide range of further medical advances and health sector transformations that could be made possible through the linkage of administrative data with large scale health data collections (such as the 'Busselton Health Study', and '45 and Up'), and private sector health insurance data.

Data that allows improved service provision

Inquiry participants flagged the potential for data relating to health service provider costs and performance, as well as de-identified linked data about health service recipients, to be used for more effective and targeted service interventions and improved health outcomes.

The New Zealand Treasury has used longitudinal data from anonymised linked administrative datasets (in this case, mental health program usage and pharmaceuticals) to identify young people at risk of poor outcomes in adulthood. By identifying a number of key characteristics that appear predictive of poor future outcomes, the analysis provided valuable insights into the effectiveness of various policies and interventions. The separation of data holdings across three levels of government and across different agencies within each of these jurisdictions, and the distrust that inhibits sharing of this data for linkage purposes, means that such analysis is not yet feasible in Australia.

Yet opportunities are emerging. The greater adoption of electronic health records in Australia (known as My Health Record) has the potential to enable more effective and holistic healthcare for patients who receive treatment from a range of healthcare providers. Pathology services are a case in point. While some duplication of diagnostic processes may be necessary for certainty or for alternative treatment plans, roughly 10% of pathology and other tests have been found to be unnecessary duplicates (CBO 2008). Using data to alert practitioners to duplicate radiology tests has been estimated to reduce the number of tests by up to 25% and test waiting time by up to 50% (Chaudhry et al. 2006), so there are substantial gains in service efficiency and patient experiences to be had from reducing duplicative effort and integrating health data.

To allow new services to emerge in response to community demand and compete with existing product offerings, potential providers need geographic information on current use of health services. For example, the Australian Dental Association highlighted that access to private health insurance data could allow for new dental practices to be established in areas of high demand.

Data that empowers individuals in managing their use of health services

Patient access to their own medical history (wherever they are, instantly) would not only improve professionals' knowledge of their patients' medical condition and reduce the number of diagnostic tests, but enable the ready and secure sharing of health information to other healthcare providers.

The chequered history of electronic health records for all Australians (beginning in 2002) now has in its most recent version, My Health Record, trial outcomes that offer genuine hope for an effective nationwide rollout.

There is substantial potential for innovation in the use of data to improve individual health. Electronic health records could incorporate and use data from monitoring devices to help to identify patients most likely to benefit from particular interventions, and predict those patients whose condition is likely to worsen (which would allow for targeted interventions by healthcare providers).

Some private sector services are already developing in Australia to allow consumers to manage their health data. Health&, for example, allows consumers to manually input and store their health data, including medical records and data from fitness devices, in a centralised location to allow better preventative health care and simpler sharing of health information between health service providers. How much more efficient and less error-prone would such transfer be if this could be done at a key-stroke? And it can, but not in Australia. That such services exist, even though they rely on manual rather than electronic input of information, is indicative of the appetite of some consumers for more control over the management of their own health data.

Risks from better data use are real but manageable

Allowing and enabling data more generally to be available and used widely would provide enormous benefits, but there are risks involved. These risks vary with the nature of the data holding, and the environment and purpose for which it is used. Public release of aggregated data on government regulatory activities, for example, may pose a very low risk of adverse consequences. Public release of data that identifies individuals who have attended a particular medical facility could, in contrast, be highly detrimental to both the individuals concerned and the reputation of the facility. Thus, the risk of harm needs to be assessed based on both the likelihood and scale of harm associated with data being more widely available. Where the adverse consequence of increased data access are considered high, the availability of the data needs to be carefully managed.

The types of risks that Inquiry participants pointed to as being most significant — related to the potential to identify persons or businesses within datasets — were:

- discrimination
- loss of control over the boundaries around the 'you' that the world sees

-
- reputational damage or embarrassment
 - identity fraud
 - other criminal misuse of the data
 - commercial harm.

That these risks exist is undeniable, but it is important not to fall victim to fear. Some, indeed most, apply to *every* form of data management, including pen and ink.

Identity theft in some form affected 126 000 Australians in 2014-15 (ABS 2016e). Most personal information used in identity theft is obtained online, either through theft, hacking or from information sent by email or placed on a website, rather than through data release or sharing. Some victims have suffered financial losses; others have reported being refused credit or accused of a crime.

Risks of identification can increase with the linkage of separate pieces of data about an individual. Matching data across individuals can also reveal more information about the activities and associations of those individuals.

These risks — and the desire for privacy and confidentiality — should not be downplayed or trivialised. They are real and important. But, many of them are able to be managed with the right policies and processes — and better managed than they are now. The likelihood of unintended or inappropriate data use needs to be carefully considered alongside the likelihood of any genuine harm or costs to the individuals or organisations concerned. Systems and processes can and should be developed to identify, assess, manage and mitigate risks related not just to data release and sharing, but also data collection and storage. This does not mean every possible data release has a technical solution: wherever and while it is not possible to reduce risks to an acceptable level, the approach being advocated by the Commission would not support public release of the data.

Giving data away

Australians give away a lot of personal information online (figure 2). For many, the information gate is (often consciously) wide open. In innumerable ways, individuals deliberately or inadvertently provide information about themselves for one purpose, which then is, or has the potential to be, used for other purposes.

- Some 68% of Australian Internet users have a social media profile, with one quarter accessing their account more than five times per day. The most popular of these sites, Facebook, soaks up information from users' computers and uses it to earn 96% of its revenue through targeted advertising. Only 12% of Internet users avoid social media for security or privacy reasons.
- Similarly, around 84% of Australians are enrolled in at least one customer loyalty program — with an average of 3.8 program memberships. While 47% recognised that a

primary reason for loyalty programs is data collection by the company, less than 2% were concerned about their privacy or felt the business knew too much about them.

- Australians have a relatively big appetite for technologies that generate or collect data (we are typically early adopters). For example, at 13% of the population, Australia has the second highest take-up rate of fitness band devices in the world. Wearable technologies, such as Fitbits, transfer data on the physical wellbeing and location of individuals back to the device provider and may be reused by it.

Some 47% of Australians report altering personal information provided online in an attempt to make themselves less identifiable (ACMA 2013a), but are often unaware that fragments of correct information on them from a wide variety of sources are being compared and matched by intelligent algorithms to form a complete and accurate picture of them.

Figure 2 **The risks that Australians take with data**



7 in 10 Australians use social media



8 in 10 participate in customer loyalty programs



5 in 10 are not fully aware of what data is collected about them and how

Only some take action to protect their privacy



5 in 10 provide false information to obtain online services



4 in 10 read privacy policies online



1 in 10 avoid social media due to privacy or security concerns

Source: Directivity et al. (2015); ACMA (2012, 2013a); OAIC (2013a)

That privacy is often said to be a concern, and yet individuals still willingly and readily hand over personal information, may seem a paradox. Because much of the data that is being generated is a byproduct of other activities and is collected in seemingly innocuous ways (mobile device location data being an example), it was once easy for individuals to dismiss it as being of secondary importance.

Today, that should not be the case. If you are using a product or service and not paying for it (or sometimes even when you are), then you are the product. This is perhaps most obvious by the ‘all or nothing’ nature of personal data requested in exchange for typically free online products and services. Often the amount of information that is required to be

provided far exceeds that necessary for delivery of the product or service that was initially sought. What you are consuming, how and when you are consuming it, is all being collected as data that is likely of considerably more value to the supplier than users may appreciate. This is not to condemn value-adding but rather to highlight the potential imbalance; and the risk to continuing consumer support that must accompany it.

Individuals typically have less choice about providing personal information to governments and may see a less immediate or personal benefit from doing so. Despite claims of a few privacy advocate groups, this Inquiry has not been presented with evidence to suggest widespread concern about the provision of personal information to governments. Indeed, the Office of the Australian Information Commissioner has found that 70% of Australians trust governments in the handling of personal information (only health service providers and financial institutions were rated higher).

Increasing data use does not necessarily increase risk

In reality, most risks of data misuse arise not through controlled sharing of data or the public release of robustly de-identified data, but rather from dataset hacking made possible by poor or outdated data collection, storage and management practices, coupled with malicious intent to gain access and use data that would otherwise not have been available.

The other avenue made possible by increased online activity is misuse of personal information that individuals have readily made public, to access other information that is not public (essentially a form of identity theft). As the value of data rises, the incentives for such exploitation rise. If data collectors fail to remain vigilant and up-to-date in technology, the present generally benign environment for collectors (public or private) may alter. Regaining social licence in such circumstances may be very difficult.

Most researchers, ethics committees, public sector data custodians and large private sector data holders are strongly motivated to handle data in ways that retain community trust and their organisation's (or own personal) reputation. But where that motivation is combined with uncertainty about privacy and secrecy requirements, or an indifference to the reservations of those contributing data, then a disproportionate reporting of bad experiences can have a long lasting chilling effect on dataset sharing and release.

That most data breaches are inconsequential and go largely unnoticed is hardly the point. The tipping point is unknown, until too late.

Tightening privacy legislation or creating new offences cannot prevent human error; nor are they likely to dissuade criminal intent. Further, to the extent that such responses inhibit well-intentioned testing of the robustness of de-identification approaches and security measures, and inflame the risk aversion of data custodians, they would represent a setback in Australia's data sharing and release efforts.

Worst of all is inaction, a feature of many failed requests for data access that were described to this Inquiry. This not only denies Australia a raft of opportunities, it also takes no account of incentives — there is a profound lack of interest amongst most researchers in government and academia in identifying particular individuals from large datasets; for them, de-identified datasets about large *groups* of people hold the answer to many pivotal questions.

Greater use of data does not mean Australians should be put at greater risk of harm. A key issue in balancing access and trust is consideration of the level of data required for different uses and how to manage well any associated risks. Near real time data that identifies individual persons or businesses carries the highest risks to privacy and security. Access to this level of data by those other than the parties to a transaction — while useful for the enforcement of some regulations (for example, traffic speed limits) and for inducing timely changes in consumer behaviour (for example, price responsive household electricity consumption) — is generally not currently necessary in order to obtain most of the benefits of data use. For analysis of market opportunities, scenario development, policy evaluation or improved delivery of many products and services, de-identified data can be sufficient, and indeed, desirable. And, of course, there is considerable data that is non-personal and non-confidential, that also needs to be made more accessible for use and re-use.

Fundamental change is needed

The legal and policy frameworks under which public and private sector data is collected, stored and used (or traded) in Australia are ad hoc and not contemporary. Privacy has carved out a space, but privacy is only one aspect of data use, and a defensive one at that. Restrictions on use for data collections in the same field, even the same institutional setting, vary significantly. Uncertainty endorses inaction.

Yet the impetus for changes in governance structures around data — changes that deal head-on with the fact that data is increasingly digital, revealing of the activities and preferences of individual people or businesses, and distributed widely in the private sector — will not diminish. It is a global movement and, to its detriment, Australia is not actively participating; and has remained nervous about making decisions.

Adding more ad hoc adjustments to existing structures and legislation will not suffice. Fundamental and systematic changes are needed to the way Australian governments, business and individuals handle data. This conclusion is based on a number of findings:

- The nature of data sources and data analytical techniques are evolving rapidly and moving away from any effective control by individuals, and will continue to do so — doing nothing is no longer an option.
- As data standards and metadata improve, digital data could be readily transferred across the economy, between sectors and across national boundaries with increasing ease. To ensure public awareness and social licence match these trends and that we make the most of them, data management frameworks need to be consistent across the economy.

-
- Incremental changes in the data management framework to date have failed to deliver a culture of making data available for widespread use. The range and volume of datasets now held in the public or private sector, that *could* potentially be made more widely available and the associated opportunities are monumental. Culture shift requires both persistent change of attitude and permission to demonstrate that.
 - There are key unanswered questions that go to the fundamental rights of individuals to control data held about them, and how individuals — as consumers — can use data more effectively for their own benefit, that lie at the heart of data availability and use. These questions necessitate an across-the-board rethink of the way data is managed.

Broad criteria shaping the recommended Framework are that it must: deliver net benefits to the community; increase the availability and usefulness of data; engender community trust and confidence in how data is managed and used; be able to adapt to higher risk; and preserve commercial incentives to collect, maintain and add value to data.

Recent progress in policy and practices around data management is acknowledged in the report: spatial data and population health have been strong for many years; linking tax and business performance data is nascent; and some State governments have been proactive. But overall, and despite positive sentiment for some time now, progress has to date been insufficient, given the broad effects possible across the social economy from the massive growth in digital data.

Thus the Commission has quite deliberately recommended the creation of a new, broad-reaching Framework that should, by design, be capable of enduring beyond current technologies, policies, personnel and institutional structures. It takes account of the significant differences in data types and associated risks and uses of each, and recognises that, while the incremental risks of making data more available might appear very small (given how much data is already in the public domain), incentives and trust nevertheless have to be maintained.

In fact, it is vital for Australia's data future that the risks of data handling are managed well. Businesses, as much as governments, rely on the willingness of the public — the source of so much of the data — to continue to trust data handling and use. Against the background of an ocean of personal data that is already public, there is now, and will be in the future, a need for continued community acceptance and trust in the handling of personal data by both governments and business. Social licence will develop if people:

- have a sound basis for believing in the integrity and accountability of entities (public and private) handling data
- feel they have some control over how their own data is used and by whom, and an inalienable ability to choose to experience some of the benefits of these uses themselves
- better understand the potential community-wide benefits of data use.

It can be difficult for a data holder to know if they have community support for use of data; but they will almost certainly know if they do not. Moreover, community acceptance of greater data use is not a one-off concept, nor is it enough that only a few better practice firms and agencies demonstrate a positive approach. Efforts of data holders to maintain community acceptance for the use of their data need to be ongoing and near universal.

Crucially, while the protections applying to personal information under the *Privacy Act 1988* (Cth) would remain, the recommended reforms would also take Australia beyond the stage of viewing data availability solely through a privacy lens. This recognises that there is much more than privacy at stake when it comes to data availability and use; this is not an Inquiry into privacy.

Although we would have preferred to find solutions that are non-regulatory, it is a clear conclusion of the Inquiry that legislative change is needed to implement the Commission's recommended reforms. This change primarily involves the creation of new Commonwealth legislation — a new Data Sharing and Release Act (DSR Act) — that would apply to all digital data.

By giving consumers new rights to use their digital data and data holders permission to be pro-active about data possibilities, the DSR Act creates a new lens through which to view data; the lens of a valuable asset being created and utilised, not merely a risk or an overhead.

The new Framework

We are recommending two facets to Australia's data Framework for the future:

1. a new right that enables both opportunities for active data use by consumers and fundamental reform in Australia's competition policy
2. a structure for data sharing and release that would allow access arrangements to be dialled up or down according to the different risks associated with different types of data, uses and use environments.

Each of these facets is discussed in detail below.

Giving individuals more control over their digital data

Australian consumers have little capacity to choose how digital data about them is used; and too often, organisations and governments make decisions (after complying with privacy principles) about the use of individuals' data on their behalf. In the face of the ubiquity of data collected, the scope to provide consumers with a greater say — within limits — on the handling of data that is sourced from them, is considerable.

The Commission is recommending that Australia’s consumers — both individuals and small and medium sized businesses (SMEs) — be afforded a new Comprehensive Right to the use of their digital data. This Right would apply to digital data holdings only. The regulation would not require businesses with paper records to digitise these in order to supply consumer data. Businesses are increasingly holding data in digital formats and it is inevitable that the value of the Comprehensive Right for consumers would grow over time. And where businesses for their own purposes do convert legacy records to digital form, such information would become potential consumer data.

While the recommended new Right for consumers provides features that match those inherent in privacy provisions (ability to view your information and propose changes to it), it is not, nor is it intended to be, a replica of privacy law. Rather, the new Right is meant to lift up the opportunity for consumers and offer a *genuine* two-way street to support their continuing willingness to supply a crucial input to business, research and public policy — namely, their data (whether obtained directly or through other channels). Consumers would no longer be just a source of data, they would rank equally with the key data collectors — businesses and governments — in being able to trade and use their data.

The ability to require the transfer of your data from one data holder to an alternative party would offer consumers the opportunity to trade safely and conveniently on their data, as business and (increasingly) governments do. To date, Australian consumers have not had much of an opportunity. Other nations (United Kingdom, European Union, for example) are more active in this regard.

Apart from building social licence through greater opportunity to use data, this Right would afford individuals and SMEs more choice about the products and services they consume, and the providers of those, and be an avenue to improve market competitiveness and innovation. No longer will it be just the collectors of data that are able to determine its uses and realise its value.

The digital data of all businesses, not-for-profits, government agencies and government business enterprises (excluding data that is collected for security purposes) should be subject to this new Comprehensive Right when an individual or SME is seeking to transfer their consumer data.

A right that will regenerate markets, and widen service choices

The right for consumers to request that their data be transferred to a third party — be it an entity in the public or private sector or a not-for-profit — is very likely to reinvigorate competition policy.

Barriers to entry created by regulation or infrastructure choke points were the key focus of the original competition policy reforms more than twenty years ago, and analysis by the Productivity Commission showed how important they were to economic growth in Australia for many years thereafter. Today’s barriers to competition include, high up on the

list, the availability of information — to consumers in order to know what is available today by way of new services; and to service providers, to identify interested consumers and improve the efficiency of operations. Information asymmetry has long been recognised in economics as a feature that weakens competitive markets.

Powerful as that may be, other benefits from consumer control of their data extend beyond competition between providers in a market. In some circumstances, the consumer may see benefits in having a copy of their data provided to an entity that is not a competitor (for example, provision of medical records to a life insurance company or provision of utility payment information to a credit provider). In other cases, it would be to form a new customer relationship, or obtain a quote that may lead to one, at the consumer's discretion. While such trading in data is readily imagined (and seen to a limited degree via comparison web sites in a few service industries) it most likely would allow consumers to exercise real choice in ways yet unimagined, and play a role in the innovation of new products and services for them, a key feature of data use in recent years.

Under the Commission's recommended plan, the scope of digital data available to be accessed and transferred under this Right would desirably be developed and agreed by participants within each industry.

The overarching outcome to be legislated is that the scope of consumer data in an industry is that which is sufficient to generate a competitive offer for an individual's custom from another provider. In other words, the type of data held on an individual or SME that a competing or complementary service provider would themselves need, and *reasonably* expect to obtain, in seeking to provide a competitive offering.

At its broadest level, consumer data should include:

- personal information (as defined in the *Privacy Act 1988* (Cth)) that is in digital form
- files posted online by the consumer
- data created from consumers' online transactions, Internet-connected activity or digital devices
- data purchased or obtained from a third party that is about the identified consumer
- other data associated with transactions or activity that is held in digital form and relevant to the transfer of data to a nominated third party.

Data that is only *imputed* by a data holder to be about a consumer — that is, data that has been created by a data holder through the application of insights or analysis such that it cannot reasonably be considered the consumer's data — should be included in consumer data *only* with industry negotiated agreement. Illustratively, if an insurer had determined through its own analysis that people who drink a lot of milk and eat red meat are very good car insurance risks compared with those who buy petrol at night and drink spirits, we would not expect that information to be included in the data to be transferred, unless an industry agreed on its inclusion.

All data does, of course, have originating sources. But it is not always reasonable to trace the trail back to its source. When multiple data sources are transformed to an extent that it is merely probable, but not certain, that a characteristic is associated with an individual consumer, this data would most likely be proprietary information of the data holder entity or perhaps a data analytics supplier. Another party seeking access to it should invest accordingly, as others have done before them.

Various terms have been suggested to describe such data that might be exempted from consumer data (or included for that matter) through industry negotiated agreement. We favour imputed data as a general descriptor — imputed being a known term in both statistics and the law. Some have suggested an exemption for value added data. This is too broad: data that has been cleansed of errors, made better through simple statistical means such as aggregated or averaged for each consumer but remains unaltered, or made machine-readable could all be construed to be value added. Consumers would be unreasonably deprived under such a descriptor. This would hardly contribute to maintaining community support in a world of increasing data exploitation.

Others have suggested that consumer data could be limited to just the transactional data generated in the direct course of the relationship between the consumer and the data custodian. In a few industries, this may be sufficient to achieve a competitive outcome for a consumer; negotiation would demonstrate that under the recommended approach. But with the exhaustive depth and breadth of data collection today and the heightened relevance of knowing your customer to achieve the outcome specified in our approach, the likelihood of this seems limited.

Similarly, there may be some data over which an entity, other than the firm that is the subject of the consumer data request, holds an intellectual property right, even where data may still be identified with a consumer. While it is legally possible to require that such data be nevertheless provided to identified consumers, this might prove to be an extreme step.

The abuse of intellectual property rights to prevent consumers accessing their data would be troubling — a sister report to this one is currently under consideration by the government. But until the Comprehensive Right is in operation and evidence of abuse, if any, of intellectual property rights emerges, the Commission prefers to allow this to be excluded from consumer data.

Industry-agreed coverage of consumer data would be determined in a data-specification process, the outcome of which would be registered with the Australian Competition and Consumer Commission (ACCC), which may approve, reject or offer interim approval.

In the absence of industry agreement as to the composition of consumer data, the broadest level definition (discussed above) should be incorporated into a consumer's request for their data held by an entity. The ACCC would determine, through the presence or absence of a registered industry data-specification what level of access a consumer was entitled to, should a dispute arise.

The joint nature of the data that is subject to the Right (that is, shared between the individual consumer and the businesses or agencies that hold the data) should ensure that incentives for data holders and collectors to collect data persist. Such entities would not be deprived of use of the asset, even as the consumer also uses it.

Transfers in action

The technological approach adopted to enable a transfer of data has received some attention.

We — and indeed, a majority of stakeholders responding to the Draft Report — are strongly against locking in a particular technological approach that all industries must adopt. In industries where frequent, real-time data transfers are needed and market participants have already made steps to enable this, transfer of data may be best achieved by the use of Application Programming Interfaces (APIs). In other industries where transfers reflect more one-off requests by consumers, alternative technology that enables the secure transfer of files may be a better approach.

Either way, standards around data formats and definitions would be necessary. We consider that participants in each industry, rather than governments, are best placed to develop these standards and determine them as part of the industry-agreed approach to transfer technology. And unlike our Draft Report, we no longer propose any additional right to opt out of data collection.

Knowing when your data has been sold

One of the most potentially pernicious practices with data is the onward trade or disclosure of data to third parties, leaving consumers unaware of who knows what about them. The damage is often not so much in monetary terms but in the feeling of exploitation. This has great capacity to undermine social licence over time, if misused. Around half of all Australians surveyed by Office of the Australian Information Commissioner (OAIC) have expressed concern about unknown organisations having obtained their personal information.

We do not propose that consumers be advised on each occasion data is traded or otherwise disclosed to a third party — the burden on businesses using contractors and outsourcing aspects of their operations could be enormous. Moreover, consumers in some areas could be inundated. But advising on which organisations data has been traded or disclosed to is a reasonable expectation of what is, after all, a joint right to data. You should surely be informed that something in which you now have a joint right is traded or disclosed to a third party.

Accordingly, entities should inform consumers about their data being traded or disclosed by including in their privacy policies, terms and conditions or on their websites, a list of

parties to whom consumer data has been traded or otherwise disclosed over the past 12 months. Such lists should be easily accessible to consumers and updated in a timely manner.

Consumers may also be at risk of loss of data access on the wind up of a firm. In such circumstances, consumers should always be advised of who now holds their data if it is transferred (as an asset) by the insolvency practitioner; or dataset owner if the data is separately sold.

Costs, timeliness and transition

We recognise that there may be costs to business associated with their adherence to the Right. There are a number of aspects of the recommendation that seek to ensure these are manageable.

First, as noted above, it is expected that industry sectors themselves would determine the scope of data to be transferred, subject to approval by the ACCC.

Second, businesses and government data holders would be able to charge for costs reasonably incurred in transferring consumer data. We fully expect that there may be a tiered approach to such charges, namely that some digital data that is of high quality, readily available, and clearly identifiable with a particular individual (such as transactions data), should be made available at low or no cost and at relatively short notice. Data stored on different (yet still digital) systems, or that is of lesser quality may require additional effort to provide in a usable format and therefore could attract a higher charge and take longer. This would be for data holders themselves to determine *and explain*.

Our intention in recommending the creation of this Right is to enhance consumer outcomes, as a contribution to sustaining community support for the role data will play in the future. Business and governments as data holders would need to adjust to this Right. Neither should have interests in creating a process that was so costly as to prohibit its take up by most if not all consumers, as this would be counter to enhancing consumer outcomes and may eventually undermine the quality of data collections.

To make the process manageable, it is surely preferable to offer the parties affected in incurring expense the chance to meet the *intent* of the Right, namely enabling consumers to use their data. This is likely to involve degrees of iteration and transition. But the clear expectation is that there would be transparency on the part of businesses and agencies. Over time as systems evolve, the time taken and the cost involved should fall as these processes become part of each firm growing its business or government agency keeping faith with its clients, and while volume of data transferred might reasonably be expected to grow.

Similarly, it is expected that businesses and government data holders themselves would likely reap benefits from system transformation and better data management, such that all of the costs would not reasonably fall to consumers availing themselves of the Right.

Support for consumers in exercising their new Right

The ACCC would be the primary government entity charged with ensuring consumers are able to transfer their data and exercise their new rights. Specifically, any charges levied by data holders for access, editing, copying and/or transferring of data should be monitored, with the methodology used by a data holder recorded, transparent (such as on the data holder's web page) and reviewable on request by the ACCC.

While recourse for consumers not satisfied with the way their new Comprehensive Right can be exercised could primarily be through the ACCC, we recognise there are other bodies — industry-specific ombudsmen, State and Territory fair trading offices, and the OAIC — that may have industry-specific skills and knowledge to deal with particular complaints. There should be a 'no wrong door' approach to this. This means the key regulators need to implement systems that enable consumer concerns to be handled with efficacy — not leave the consumer straddling a regulator abyss.

While the changes proposed aim to enable consumers to exercise more control over the collection and use of their data, the onus remains on individuals to make responsible choices regarding to whom they provide personal information in the first instance and for what purposes.

Comprehensive credit reporting

In some circumstances, collating consumer data may offer net public benefit in making markets more efficient. A specific case is covered in the terms of reference for this Inquiry: Comprehensive Credit Reporting (CCR).

The Productivity Commission has previously found comprehensive credit reporting to be desirable and, consistent with the approach of New Zealand, the United Kingdom and the United States, a voluntary approach to data input should continue to be pursued, unless it is clear that a critical mass of accounts is not achievable on that basis. In the event that voluntary participation in the scheme remains below a critical mass of 40% in mid 2017, the Australian Treasury should proceed with developing draft legislation to mandate comprehensive credit reporting.

We note this is a date almost upon us, but those who have argued for more time could do their case greater service by noting that the industry has had since December 2015 to show clear movement, and a substantial notice period of CCR's imminence prior to that. Greater movement, if it occurs by December 2017, may convince the government to hold off proceeding to parliament. But the preparation of legislation to make participation mandatory cannot be a shock after such a time; and those with subsidiaries in New Zealand must surely have gained relevant experience in CCR participation from there.

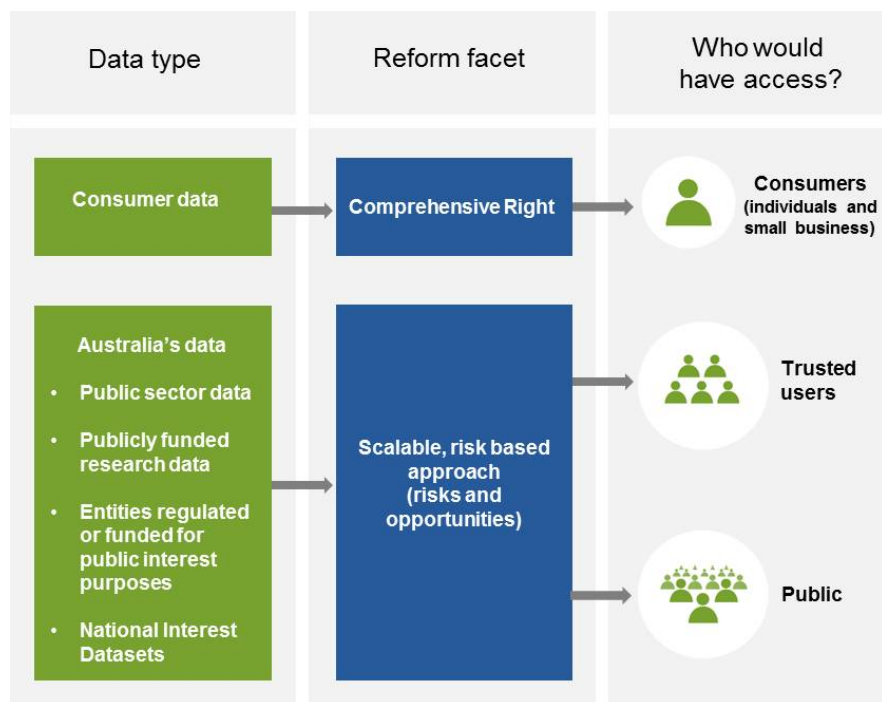
A structure to give substance to data aspirations

At both Commonwealth and State/Territory levels of government, there is an aspiration to improve the internal and external release and sharing of their data, with the objectives of creating opportunity for innovation by firms, lifting accountability and improving service delivery. New South Wales and South Australia have created legislative frameworks to put substance into aspiration. The Commonwealth has yet to do so.

Thus a substantive element of our reform Framework is a structure that provides institutional capability to seriously address these objectives and remove impediments to a consistent approach to handling data with varying levels of value and risk (figure 3).

Absent legislated permission to utilise the opportunity presented today by data analytics, aspiration will remain primarily a hope carried by well-intentioned but often under-equipped individuals.

Figure 3 **Framework of the recommended approach**



Institutions

In addition to focussing on institutional capability, our approach seeks to clearly define institutional accountability. The Commission has recommended both central agency responsibility for data in each jurisdiction and establishment of a new statutory office holder, the National Data Custodian (NDC). The NDC would be established as a statutory officer with a small advisory board to provide leadership and ongoing updating of the new

data Framework, including managing the broader ethical considerations around data use that are increasingly arising with the mass digitisation of data (box 2).

One of the NDC's primary activities authorised in the new DSR Act would be to accredit the processes and capabilities of a suite of Accredited Release Authorities (ARAs). The NDC would also publish guidance on data use — both opportunity and threat — for the benefit of ARAs and other data custodians, and to update continuously its accreditation processes. It would audit ARAs from time to time for compliance with both guidance material and the terms of Commonwealth funding.

ARAs would be sectoral hubs of expertise, independent entities that are funded by the Commonwealth (but may be based in a State or Territory, or even a not-for-profit entity that has a public interest focus). They would be tasked with:

- developing and integrating datasets drawn from across a sector (and jurisdictions) with high prospect to improve data use and release
- assisting a field of data custodians to improve the curation and quality of datasets to be released (including de-identifying or linking where necessary)
- facilitating timely updates and ongoing dataset maintenance
- approving trusted users of more sensitive data
- determining whether a dataset that they are responsible for should be shared or released.

Under this structure, access arrangements and safeguards could be dialled up or down by the ARA according to the nature of opportunities and risks likely. ARAs would have the capability necessary to make these judgements and apply them effectively. The extent to which use of a particular dataset(s) could be enabled to provide broad benefits beyond those currently derived would be a primary factor in the Commonwealth choosing to fund an ARA.

ARAs would, under governance arrangements agreed with cooperating dataset custodians, be responsible for whether a dataset is available for public release or limited to sharing with trusted users. Dataset custodians would retain their legislated responsibility for original data contributed to ARAs and so make the choice regarding its release or sharing in that *original* form; but where an ARA goes on to transform data (for example via integration, linkage, de-identification or use of Artificial Intelligence) the *transformed* dataset would become the primary responsibility of the ARA. The ongoing maintenance and updating of the transformed dataset would necessitate cooperation between ARAs and custodians of component datasets.

Box 2 Key participants in the new Framework

The National Data Custodian (NDC)

The new position of the NDC creates a role that would provide the focus of improved national effort to lift access to data and use of data analytics to contemporary standards and objectives. The NDC would have responsibility for enabling effective use of data; oversight, guidance and updating operations of the national data system; and be instrumental in assessing for designation possible National Interest Datasets (NIDs). The NDC would also accredit release authorities within the reformed national data system and have broad responsibility for determining which release authority is most suitable to handle the ongoing management of key datasets. The Office of the NDC would include a dedicated ethics adviser capable of providing advice and guidance on ethical issues related to data access and use. A small advisory board would support the NDC in technical matters of data access and use.

Accredited Release Authorities (ARAs)

ARAs would largely be existing public sector agencies (Australian Government or State/Territory government) that already release data but would now be funded to take on additional responsibilities as an ARA (the Australian Institute of Health and Welfare is a working model for an ARA).

ARAs would play an important role in deciding whether a dataset is available for public release or limited sharing with trusted users, approving trusted users, curating datasets and assisting dataset custodians with curation and the development of metadata, ensuring the timely update and maintenance of datasets, and supporting the linkage of NIDs and other datasets. Given the emphasis on sectoral expertise, these entities would have a long track record of trusted data management in their particular areas of focus. It is envisaged that ARAs would also perform an important advisory role on technical matters, to the NDC, government, and to the broader community of data custodians.

Trusted users

Trusted users would be accredited by ARAs to access data under its control or governance. They may include any individual working in an entity that has in place the necessary data governance requirements to safely handle the datasets in question and a signed legal undertaking that sets out safeguards for use and recognises relevant privacy requirements. Personnel in relevant Commonwealth and State/Territory government agencies and publicly funded research bodies should be eligible for accreditation as a trusted user. Private researchers should be able to achieve trusted user status, once their employer's commitment to the ARA's standards and approval requirements is evident.

Those wanting access to datasets maintained by ARAs would require the necessary governance structures and processes to address the risks associated with data use or release as a consequence of research findings, including access to secure computing infrastructure. Accreditation as a trusted user should enable a researcher to access a particular dataset or datasets on an ongoing basis for a program of work, rather than project by project.

There may remain some trusted users who access datasets held by data custodians under a variety of conditions set by each custodian. The introduction of ARAs would not alter these arrangements.

High value datasets

The Commission's Terms of Reference asked for advice on selecting high value datasets. The range of possible nominations for this is very broad and many submissions offered candidates. We have not, however, nominated any winners. Instead, informed by the breadth of advice and the impediments to maximising net benefit for Australians from further public investment in our data, we have designed processes to be applied to deliver the kind of outcome envisaged by recent governments' statements and our own assessment.

National Interest Datasets

Governments across Australia hold enormous amounts of data, but mostly lag behind other comparable economies in beneficially using data beyond the purposes for which it was initially collected, or allowing others access to do so. Australia's private sector data holders are often more innovative in their use of data. The public sector needs to spend the time and effort to identify strong candidates for the kind of investment via data analytics that the private sector is now applying. Absent this effort, the public sector will be denuded of core analytical capabilities in a digitally-driven future, and forgo the opportunities afforded by those capabilities.

The first step in this effort is to identify National Interest Datasets (NIDs).

From submissions in response to the Draft Report, it is evident that the concept is very positively viewed, albeit with uncertainty about the methodology of making the choices amongst many obviously very important datasets.

The simple summary of the process outlined below (and in more detail in the Report), concentrates on additionality. What additional community-wide benefit could be achieved by ensuring and consistently enabling wider sharing, use and possibly release of that data?

Designation as an NID would achieve wider sharing, use, and where appropriate, release by:

- providing sufficient and predictable funding for data curation and updating
- aggregating jurisdictionally-separated datasets
- providing a framework in which to better link datasets across sectors and fields of endeavour.

For instance, significant improvements could come from aggregating data across the States and Territories in health, education, social welfare, child support, aged care, and better linking them with elements of datasets from other fields — the population census, taxation, employment, business ownership, telecommunications, private health insurance or housing. This is not an exhaustive list. State datasets in justice, infrastructure, land use and

property ownership could be similar candidates for permanent linkages and curation by ARAs with a plan for delivering identified national benefits from this investment.

Such benefits would be identified on an ongoing basis by the NDC, after an initial set of NIDs is put forward with the DSR Bill.

Designating datasets as national interest collections would also signify their value as resources collected in the national interest. But this is not about labelling a collection as important in principle; its purpose must be additional national benefit in practice.

NIDs are thus likely to have the following characteristics:

- relatively few in number, certainly not in the thousands
- linked, integrated, transformed (for example, by de-identification or use of Artificial Intelligence) to suit a prior determined scope of outcomes
- offering clearly described public interest benefits of national application
- have been confirmed by public review (we are proposing consideration by a parliamentary committee).

For datasets designated as NIDs, new access and use arrangements devised under the DSR Act would take precedence over existing restrictions — Commonwealth or State — to access.

Impediments from jurisdictionally inconsistent access arrangements, confidentiality standards, collection purposes, obligations to seek approval for use and privacy rules would be replaced by a single legislated modern fit-for-purpose regime to apply when data is designated an NID and transferred to an ARA.

The extent to which this modern fit-for-purpose regime transforms each NID would be determined ultimately by the final legislated design and the restrictions in the head legislation of a data collection that is to be incorporated into an NID.

The DSR Act should draw upon all Commonwealth Constitutional powers — including of particular value, the telecommunications powers. This is likely to be effective with many datasets. Nevertheless, advice to us is that depending on the specific chosen datasets, some collection-specific regulation may also be required.

The intent is nevertheless that by the act of bringing forward datasets and arguing in public fora for their freedom to be used in the national interest, both the public and private sector (the latter in rare circumstances, exemplified by private education and health data) can create an asset and governance structure that each would be comfortable with, and which the Commonwealth could support in order to deliver the identified benefits.

We have recommended a parliamentary committee process be used to expose the NDC's analysis of national benefit to public comment. It is not often that the Executive willingly offers parliament the opportunity to critique its proposals but we see a number of benefits to this process.

First, the term national interest is an essentially arbitrary one. Rather than attempt to pin it down by definitional drafting, we propose that the comment of our national representatives is used. Second, the approach is consistent with seeking wherever possible to obtain social licence via public involvement. Third, it would allow for the NDC along with others in the community to continuously make known why better use of data is indeed in the national interest. This educational approach to community consultation would be far more effective than what is often simply box-ticking consultation. In other words, we recommend engaging in a full parliamentary scrutiny process prior to designating NIDs by disallowable instrument.

Datasets afforded NID status should be maintained as a national asset for a period of at least 10 years.

A listing of all NIDs that have been publicly released or are potentially available to share, the relevant ARA for that dataset and custodians of component datasets, and a contact point, would be included on a central website, such as data.gov.au. This would enable potential users of these to know of the dataset's existence and how to gain access to it.

Datasets that provide additional opportunities for Australia-wide gains

Most datasets would not be NIDs. Beyond them, the ARA model is also intended to offer the opportunity for the much larger range of datasets in public hands to be better curated and managed for cooperative release by, or with the assistance of, ARAs. Thus, any dataset provided to an ARA under the DSR Act would also be subject to the single modernised legislative regime to enable improved data access. And Commonwealth data custodians, whether ARAs or not, should observe and apply the NDC's guidance on better data practice; and report on it annually.

Fee for service arrangements, such as that taken by the Australian Institute for Health and Welfare today, would support the resourcing required by ARAs.

The special case of higher risk data

Some data that identifies (or has the potential to identify) individual persons or businesses is already shared in a very limited way with policy developers and with researchers within government and/or the research community. This data is typically used for targeted program and product/service delivery, for research (such as rare medical conditions) where there are very small populations involved, and (in some limited cases) for regulatory compliance.

The current process requires, however, multiple approvals, delay and (often) work-arounds for all involved, even where there is great acceptance of the validity of the work. It surely can be improved, safely.

Depending on the particular dataset, access requests (even from within the same government) can require separate and duplicative agreement of multiple dataset owners, custodians and stewards, integration units, ethics committees, other advisory bodies, and the individuals about whom the information was collected. Each guideline and approval step may be reasonable in its own right, but collectively it is costly and can be self-defeating. We were advised of Australian researchers abandoning attempts to access Australian data in favour of UK, EU or US data.

The Commission recommends streamlining access to identifiable data within and between Australian governments, and for the limited range of other trusted users with which such data is shared. Researchers in fields that require access to identified data already have strong professional reasons to avoid misusing it.

Simplified but still protected access to identifiable data would be enabled via ARAs. This would include all requests for access to identified data within ARA-managed datasets — whether NIDs or otherwise (the latter being where original data custodians have offered data to an ARA, thus allowing an ARA to create a transformed ARA-held dataset). Governance arrangements and technical capability to enable such access with safeguards would form a core part of the NDC’s accreditation for that ARA.

Current duplicative efforts of ethics committees and lack of recognition of approvals granted should be addressed to streamline access to data by researchers. We are recommending reforms to both the registration processes for human research ethics committees and approval processes.

The capacity to use information without obtaining consent of individuals would be extended under the DSR Act to cover all public interest research, rather than be limited to health and medical research purposes as is the case at present. The NDC would be charged with resolving whether research in question was public interest or not.

Access would occur in a specified secure computing environment with output from the dataset reviewed (by an automated process, where possible) prior to project completion to ensure risk management approaches had been satisfactorily implemented.

Responsibility for appropriate use of datasets would rest with trusted users and their institutions, with incentives to maintain necessary safeguards, including assurance arrangements from institutions involved, up to and including financial bonds where the DSR Act may not have applicability (such as in the case of access by foreign institutions).

Making other publicly funded data readily available to all

Government understanding of the demand for data and its value is limited. Novel concepts and applications for data are arising continuously and anticipating how they might benefit the wider public interest is a continuous learning process.

While there may at times be good reasons for governments' inability to derive value from its data holdings — governments are not entrepreneurial nor would we necessarily want them to be — risk aversion is not desirable where it results in the public interest being poorly served.

There needs to be a shift in emphasis from only releasing data on request for particular projects, toward actively pushing data out in a coordinated way. In principle, all non-sensitive datasets in fields where there are burgeoning opportunities and capability would be opened up and released, as resources and sectoral demand allow.

This includes data that, while it may identify individuals or businesses, is already in the public domain in some form (property ownership, for example). A realistic assessment of the risks associated with public release of identifiable information that is already public in a less accessible form, should be undertaken. It also includes data that may identify and be used to evaluate the performance of publicly funded or regulated services.

Such an approach has the potential to make a marked difference to the range and volume of data available for decision making, innovative activity and improved service delivery in the community.

The challenges in achieving this should not be underestimated. There is a very real culture of risk aversion and risk avoidance in the public sector when it comes to data release.

Not just an academic exercise

There is a need for the research community to also put its house in order when it comes to data sharing. Just as government data custodians should consider that they hold data not solely for their own purposes but in the public interest on behalf of citizens, so too should the data of publicly funded research be available beyond the initial researchers. And where it is not, much better justification and record keeping is needed, to at least enable other researchers to learn what data has already been collected.

Leadership

New arrangements for data access will require, within Governments, strong and consistent leadership, from the Ministerial level as well as the upper echelons of the bureaucracy.

The application of the Framework and DSR Act will offer a clear sign of changed attitudes and permission to take a proactive approach to data use, but the first misstep will be a test. Custodians will not change culture without steady leadership in the face of short-term adversity. We are working to change attitudes and processes that have been entrenched for decades.

While release of public sector data would be the focus of governments, it is anticipated that once governments start to more actively push data out, this will encourage private entities to do likewise and to profit from doing so. That is, across the economy the value will shift from being embodied in the data itself, to being derived from the clever analysis and use of data.

Implementation

Establishment of the full Framework in this Report requires a firm implementation plan; and one that at the outset envisages implementation of all elements.

An underlying goal of the implementation plan should be to move forward in a way that retains and ultimately builds community confidence in what will be an all-encompassing data exchange between consumers, governments, researchers and business. While this Inquiry has created a new level of awareness of data opportunities amongst data holders and regulators, change of this order cannot be launched into the public arena without further planning and effective communication of intent and actions.

State and Territory agencies are an important part of the new Framework. It is essential that they be included in determining the details of its implementation — they would be beneficiaries of wider data accessibility and if they are allowed to realise these benefits, then would more likely also be contributors of datasets. They also hold much of the country's skills and expertise in some nationally-relevant sectoral data.

Apart from the new functions created — the Office of the National Data Custodian, a parliamentary committee (new or existing) for scrutiny of National Interest Datasets, and Accredited Release Authorities — the recommended reforms involve additional roles for several existing bodies. The ACCC would have additional functions in implementation of the Comprehensive Right for consumers; agencies with expertise in data integration and release, such as the Australian Institute of Health and Welfare, the Australian Bureau of Statistics, CSIRO Data61, and State linkage bodies, would have additional functions (either funded or charged) in implementation of the risk-based structure for data sharing and release.

The Australian Government should set an ambitious timeline for reform implementation, and move quickly where administrative change is all that is required to put substance into new policies (figure 4). We envisage that some of the recommendations can, and should, be implemented very quickly:

- allowing State-based linkage units to link Commonwealth data;
- abandoning the current policy obligation to destroy datasets and linkage keys on completion of research;
- publication of data registers for all public and publicly funded data holdings;
- administrative appointment of the National Data Custodian; and
- commencing the drafting and consultation processes for the new DSR Act.

To delay these would create a debilitating loss of policy momentum and forgo the possibility of early gains in community acceptance for reforms.

Beyond that *but still in the short term*: appointment of an advisory board and ethics advisor to the Office of the National Data Custodian; reform of human research ethics committees processes around registration and approval recognition; development of processes for the accreditation of ARAs and nomination of National Interest Datasets; and establishment of an early set of priority projects to reach agreement on industry standards for transfer of consumer data under the Comprehensive Right; should be pursued relatively quickly.

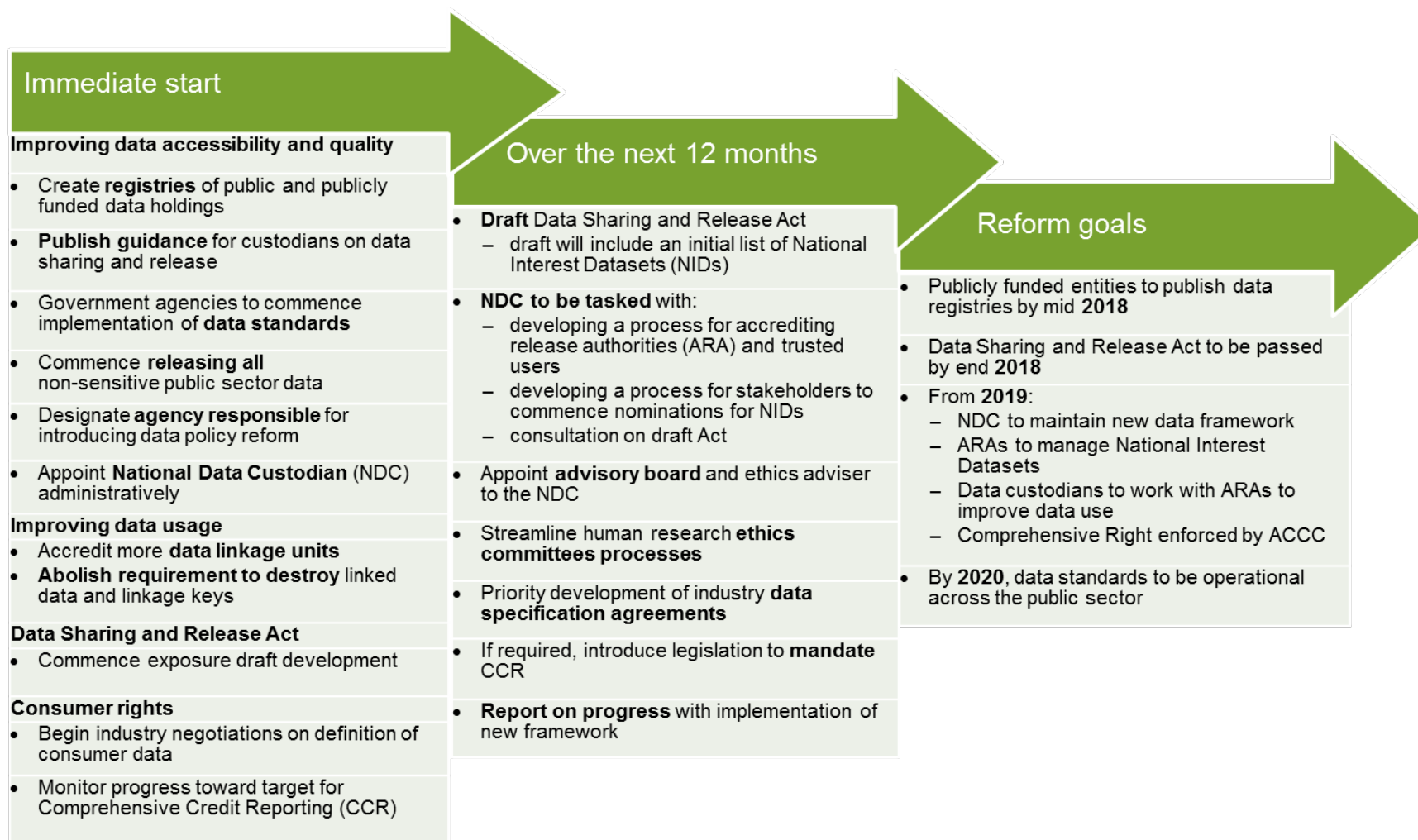
The new data Framework is intended to alter approaches to data holdings across public and private sector entities, including research bodies and not-for-profits. As noted earlier, we see no basis for distinguishing between these entities in their attitudes to ensuring data is more accessible.

The Commission cautions against any temptation to decouple parts of the reform Framework. In a project that aims to create new opportunity for both public and private benefit, each element supports the others. While gains from reforming public sector data access are likely to be substantial and benefit the public over time, governments tempted to ignore consumers' data rights, for example, would make their life difficult in trying to get the social licence needed for data reforms more generally.

It is only by allowing individuals and SMEs as consumers to share benefits of greater data use directly that governments and business would garner the community approval necessary for ongoing data collection and use.

Other developed countries have been making long term commitments necessary to engender community acceptance and reap benefits from the opportunities that mass digitisation of data now presents. It is time for Australia to reach further.

Figure 4 Implementation timeline for the Commission’s key recommendations



Findings and recommendations

The current data environment

FINDING 1.1

Australia's provision of open access to public sector data is below comparable countries with similar governance structures, including the United States, the United Kingdom and New Zealand.

While there remains considerable scope to improve the range of datasets published (and, correspondingly, the diversity of agencies and research bodies publicly releasing data), poor formatting and the lack of frequency with which data is publicly updated are reducing data usability.

FINDING 2.1

The benefits from greater access to data would be widespread, but consumers, in particular, have much to gain, collectively, from action on Australia's data sharing and release arrangements.

FINDING 3.1

Individuals are likely to be more willing to allow data about themselves to be used by private and public organisations, provided they understand why and how the data is being used, can see tangible benefits, and have control over who the data is shared with.

FINDING 3.2

A wide range of more than 500 secrecy provisions in Commonwealth legislation plus other policies and guidelines impose considerable limitations on the availability and use of identifiable data. While some may remain valid, they are rarely reviewed or modified. Many would no longer be fit for purpose.

Incremental change to data management frameworks is unlikely to be effective or timely, given the proliferation of these restrictions.

FINDING 3.3

Data integration in some jurisdictions (particularly Western Australia and New South Wales) has progressed in some fields, but highlights a lack of action in equivalent fields at both Commonwealth and State level, and reveals the large unmet potential in data integration opportunities.

FINDING 3.4

The boundaries of personal information are constantly shifting in response to technological advances and new digital products, along with community expectations.

The legal definition of personal information, contained in the *Privacy Act 1988* (Cth), has always had an element of uncertainty, and is managed by guidelines. In the face of rapid changes in sources and types of data, outcome-focused data definitions remain essential. But practical guidance (that data custodians and users can rely on) is required on what sorts of data are covered by the definitions.

FINDING 3.5

Despite recent statements in favour of greater openness, many areas of Australia's public sector continue to exhibit a reluctance to share or release data.

The entrenched culture of risk aversion, reinforced by a range of policy requirements and approval processes, and often perverse incentives, greatly inhibits data discovery, analysis and use.

The lack of public release and data sharing between government entities has contributed to fragmentation and duplication of data collection activities. This not only wastes public and private sector resources but also places a larger than necessary reporting burden on individuals and organisations.

FINDING 3.6

Large volumes of identifiable information are already published online by individuals or collected by various organisations, with or without explicit consent.

Breaches of personal data, often compounded by individuals' unwary approach to offering data, are largely dominated by malicious database hacking or criminal activity. By comparison, breaches due to sharing or release are rare.

A new Framework for sharing and release

FINDING 4.1

Comprehensive reform of Australia's data infrastructure is needed to signal that permission is granted for active data sharing and release and that data infrastructure and assets are a priority. Reforms should be underpinned by:

- clear and consistent leadership
- transparency and accountability for release and risk management
- reformed policies and legislation
- institutional change.

FINDING 4.2

Community trust and acceptance will be vital for the implementation of any reforms to Australia's data infrastructure. These can be built through enhancement of consumer rights, genuine safeguards, transparency, and effective management of risk.

A Comprehensive Right for Consumers

RECOMMENDATION 5.1

Consumer data must be provided on request to consumers or directly to a designated third party in order to exercise a number of rights, summarised as the Comprehensive Right to access and use digital data. This Comprehensive Right would enable consumers to:

- share in perpetuity joint access to and use of their consumer data with the data holder
- receive a copy of their consumer data
- request edits or corrections to it for reasons of accuracy
- be informed of the trade or other disclosure of consumer data to third parties
- direct data holders to transfer data in machine-readable form, either to the individual or to a nominated third party.

Where a transfer is requested outside of an industry (such as from a medical service provider to an insurance provider) and the agreed scope of consumer data is different in the source industry and the destination industry, the scope that applies would be that of the data sender.

RECOMMENDATION 5.2

The Australian Government should introduce an outcome-based definition of consumer data that is, as an overarching objective, data that is sufficient to enable the provision of a competing or complementary service or product for a consumer.

In the relevant service or product context, consumer data is digital data, provided in machine-readable format, that is:

- held by a product or service provider, and
- identified with a consumer, and
- associated with a product or service provided to that consumer.

Participants in an industry should determine the scope of consumer data relevant to their industry (where an industry in this context would be determined by a broad description of the service). This should be in the form of a data-specification agreement.

Data-specification agreements should also articulate: transfer mechanisms, and security of data, to ensure that data use is practical and robust to technology updates; and the requirements necessary to authenticate a consumer request prior to any transfer.

These agreements should be registered with the ACCC, which may offer interim approval where an agreement has been reached but other industry agreements have been prioritised for approval.

In the absence of such agreement, consumer data must be in machine-readable form and include all of:

- personal information, as defined in the *Privacy Act 1988* (Cth), that is in digital form
- information posted online by the consumer
- data created from consumers' online transactions, Internet-connected activity, or digital devices
- data purchased or obtained from a third party that is about the identified consumer
- other data associated with transactions or activity that is relevant to the transfer of data to a nominated third party.

Data that is solely imputed by a data holder to be about a consumer may only be included with industry-negotiated agreement. Data that is collected for security purposes or is subject to intellectual property rights would be excluded from consumer data.

A consumer for the purposes of consumer data should include a natural person and an ABN holder with a turnover of less than \$3m pa in the most recent financial year.

Data that is not able to be re-identified to a consumer in the normal course of business within a data holder should not be considered consumer data.

The definition should be included in a new Act for data sharing and release (Recommendation 8.1). Given the need for consumer data to have broad applicability, the outer boundary definition and reference to ACCC registered industry-specific definitions should also be included within the *Acts Interpretation Act 1901* (Cth). Consequential amendments to other legislation in the future would ensure harmonisation across federal laws.

RECOMMENDATION 5.3

All holders of consumer data should include in their privacy policies, terms and conditions, or on their websites a list of parties to whom consumer data has been traded or otherwise disclosed over the past 12 months.

On the windup of an entity that holds consumer data, consumers should be informed if data to which they hold a joint right has been traded or transferred to another entity. For businesses entering formal insolvency processes, insolvency practitioners should ensure consumers have been informed. For businesses closing but not in insolvency proceedings, the entity acquiring consumer data should inform consumers of this fact and give them the opportunity for data collection to cease.

RECOMMENDATION 5.4

The Australian Government should provide for broad oversight and complaints handling functions relating to the use of the Comprehensive Right. Accordingly, the Australian Competition and Consumer Commission (ACCC) should be resourced to undertake the following additional responsibilities:

- approving and registering industry data-specification agreements and standards
- handling complaints in relation to a data holder's failure to meet the terms of the Comprehensive Right, including in regard to the scope of consumer data
- educating consumers (in conjunction with State And Territory fair trading offices) on their rights and responsibilities under the Comprehensive Right
- assessing the validity, when requested or at their discretion, of charges levied by data holders for application of the Comprehensive Right.

The Office of the Australian Information Commissioner and industry ombudsmen should, in order to ensure a 'no wrong door' approach to handling consumer engagement, coordinate with the ACCC on the receipt and handling of consumer complaints on data access and use.

RECOMMENDATION 5.5

The Australian Government should adopt a minimum target for voluntary participation in Comprehensive Credit Reporting of 40% of all active credit accounts, provided by Australian Securities and Investments Commission (ASIC)-licensed credit providers, for which comprehensive data is supplied to the credit bureaux in public mode.

If this target is not achieved by 30 June 2017, the Government should circulate draft legislation by 31 December 2017, to impose mandatory participation in Comprehensive Credit Reporting (including the reporting of repayment history) by ASIC-licensed credit providers in 2018.

The Office of the Australian Information Commissioner and ASIC should consult with other regulators, industry groups and consumer advocates to collaboratively consider whether there is a need for a hardship flag in credit reporting.

The Department of the Treasury should be given responsibility for monitoring and publicly reporting on a regular basis on participation in Comprehensive Credit Reporting.

A risk-based approach to data sharing and release

RECOMMENDATION 6.1

As an immediate objective, all Australian governments should direct the early release of all non-sensitive publicly funded datasets — whether held by a government agency or other body receiving public funding for data collection activities.

A realistic assessment of the risks attached to public release of identifiable information that is already public (in a less accessible form) should be undertaken by all governments, with the intention of releasing low risk data, and mitigating risks where possible to enable far greater public release of data, including that which could be used for program or agency performance management purposes.

Agencies should report annually on the proportions of their datasets made publicly available, shared, and not available for release.

RECOMMENDATION 6.2

Additional qualified entities should be accredited to undertake data linkage.

State-based data linkage units should be able to apply for accreditation by the National Data Custodian (Recommendation 6.6) to allow them to link Australian Government data.

RECOMMENDATION 6.3

All Australian governments entering into contracts with the private sector that involve the creation of datasets in the course of delivering public services should assess the strategic significance and public interest value of the data as part of the contracting process.

Where data is assessed to be valuable, governments should retain the right to access or purchase that data in machine-readable form and to subsequently apply any analysis and release strategy that is in the public interest.

The Australian Government Department of Finance should modify template contracts to, by default, vest access and purchase rights in governments, and avoid the need for negotiating separate rights in each contract. State and Territory governments should adopt a similar approach.

RECOMMENDATION 6.4

Publicly funded entities, including all Australian Government agencies, should create comprehensive, easy to access registers of data, including metadata and linked datasets, that they fund or hold. These registers should be published on data.gov.au. Where datasets are held or funded but are not available for access or release, the register should indicate this and the reasons why this is so.

States and Territories should create an equivalent model for their agencies where such registers do not exist. These should, in turn, be linked to data.gov.au.

A reasonable timeframe in which to achieve this is within one year (by March 2018).

RECOMMENDATION 6.5

In determining datasets for public release, a central government agency in each jurisdiction with overarching policy responsibility for data should offer a public process whereby datasets or combinations of datasets can be nominated, with a public interest case made, for release.

A list of requested datasets, and decisions regarding dataset release or otherwise, should be transparent and published online — in the Commonwealth's case, on data.gov.au.

RECOMMENDATION 6.6

The Australian Government should establish an Office of the National Data Custodian (NDC) to take overall responsibility for the implementation of data management policy, in consultation with all levels of Government.

The Office of the NDC should have responsibility for:

- broad oversight and ongoing monitoring of and public reporting on Australia's national data system and the operation of the new Data Sharing and Release Act (recommendation 8.1)
- preliminary assessments for, and recommending designation of, National Interest Datasets (recommendation 7.1)
- accrediting release authorities, be party to determining a funding agreement for Accredited Release Authority (ARA) activities, and promoting cooperation between ARAs
- managing complaints about ARA processes
- providing practical guidance material for ARAs and data custodians on matters such as risk management, data curation and metadata, data security, data de-identification and trusted user models
- advising on ethics and emerging risks and opportunities in data use.

The Office of the NDC should include a small advisory board, comprising members with technical skills related to the NDC's activities, and a dedicated ethics adviser.

The NDC role should be filled administratively by the end of 2017 to be operational by the time that new draft legislation for data access is completed for public consultation (Recommendation 10.2).

RECOMMENDATION 6.7

The National Data Custodian should streamline approval processes for access to data by:

- issuing clear guidance to all Australian Government data custodians on their rights and responsibilities, ensuring that requests for access to data they hold are dealt with in a timely and efficient manner and are consistent with the risk management approach to be adopted by Accredited Release Authorities (ARAs)
- requiring that these data custodians report annually on their handling of requests for data access, including requests from ARAs.

State and Territory governments may opt in to these approaches to enable use of data for jurisdictional comparisons and cross-jurisdictional research.

RECOMMENDATION 6.8

Selected public sector and public interest entities should be accredited as release authorities. Accreditation should be determined based on sectoral expertise, capability, governance structures, and include consultation throughout the relevant sector.

Accredited Release Authorities (ARAs) would be responsible for:

- deciding (in consultation with original data custodians) whether a dataset is available for public release or limited sharing with trusted users
- collating, curating, linking and ensuring the timely updating of National Interest Datasets and other datasets
- offering advice, services and assistance on matters such as dataset curation, de-identification and linking
- providing risk-based access to trusted users.

ARAs should be fully operational from the beginning of 2019.

RECOMMENDATION 6.9

All Accredited Release Authorities must have and publish formal risk management processes to effectively assess and manage the risks associated with sharing and release of data under their control.

Standardised, access-friendly Data Sharing Agreements should be implemented with external data providers and users to formalise the activities that can take place with identifiable and de-identified data.

Risk management processes should be regularly reviewed and revised to account for new and emerging risks.

RECOMMENDATION 6.10

Funding of Accredited Release Authorities (ARAs), for the purposes of data management, curation, storage and access should be set via a funding agreement with the National Data Custodian.

ARAs should have the power to charge fees sufficient to recoup costs where ARAs undertake requested work beyond that envisaged in their funding arrangement with the National Data Custodian.

In assessing the scope to undertake such activities, ARAs must ensure they do not detract from their primary focus on the public benefits of enabling greater access to, and use of, data (which is the basis for their accreditation and funding).

RECOMMENDATION 6.11

The Office of the National Data Custodian should be afforded the power to require an audit of a data custodian's de-identification processes and issue assurance of de-identification practices used.

RECOMMENDATION 6.12

Accredited Release Authorities (ARAs) should be given responsibility to grant, on a continuing program-wide basis, data access to trusted users from a range of potential entities that:

- have the necessary governance structures and processes in place to address the risks of inappropriate data use associated with particular datasets, including access to secure computing infrastructure, and
- have a signed legal undertaking that sets out safeguards for data use and recognises relevant privacy requirements.

In assessing trusted user access, the ARAs should accept existing current approvals of the trusted user's work environment.

Trusted user status for use of identifiable data would cease for that user when they leave the approved environment, when a program is completed, or if a data breach or mishandling occurs in that same environment and/or program.

RECOMMENDATION 6.13

Accredited Release Authorities (ARAs) and data custodians should be required to refer suspected and actual violations of data use conditions that have system-wide implications to the National Data Custodian.

Clarification should be issued detailing how this process would interact with the *Privacy Amendment (Notifiable Data Breaches) Act 2017 (Cth)*.

RECOMMENDATION 6.14

Progress by individual research institutions receiving Australian Government funding in making their unique research data and metadata widely available to others should be openly published by those institutions, with reference to past performance.

All bodies channelling public funds for research, such as the National Health and Medical Research Council and Australian Research Council, should similarly require in future funding agreements with research applicants that data and metadata is to be publicly available, and publish the results of progress on this for their funded projects.

On completion of projects, research institutions should include in their reports details of when and how other researchers can access the project's data and metadata.

RECOMMENDATION 6.15

Processes for obtaining approval from human research ethics committees (HRECs) should be streamlined.

To achieve this in the health sector:

- All HRECs should be required to register with the National Health and Medical Research Council (NHMRC). The NHMRC should receive funding to expand its current registration process, to include audits of registered HRECs.
- To maintain their registration, HRECs must implement efficient and timely approval processes, which ensure projects are not unduly delayed. The time taken to consider and review projects should be reported to the NHMRC, and included in the annual report on HREC activity.
- As a condition for registration, all HRECs and the institutions they operate in would be required to accept approvals issued by certified HRECs for multi-site projects, without additional reviews. The Australian Health Ethics Committee should develop uniform review processes to be used by certified HRECs.
- The Council of Australian Governments' Health Council should sign an intergovernmental agreement that extends the existing National Mutual Acceptance Scheme to all jurisdictions, including the Commonwealth, and all types of projects. As part of this agreement, all jurisdictions should also implement streamlined governance approvals.

RECOMMENDATION 6.16

The *Privacy Act 1988* (Cth) exceptions that allow access to identifiable information for the purposes of health and medical research without seeking individuals' agreement, should be expanded in the legislative package that implements these reforms to apply to all research that is determined by the National Data Custodian to be in the public interest.

RECOMMENDATION 6.17

The Australian Government should abolish its requirement to destroy linked datasets and statistical linkage keys at the completion of researchers' data integration projects. Where an Accredited Release Authority is undertaking multiple linkage projects, it should work towards creating enduring linkage systems to increase the efficiency of linkage processes.

Data custodians should be advised as part of early implementation of this reform package to use a risk-based approach to determine how to enable ongoing use of linked datasets. The value added to original datasets by researchers should be retained and made available to other dataset users.

National Interest Datasets

RECOMMENDATION 7.1

The Australian Government, in consultation with State and Territory governments, should establish a process whereby public (and in some exceptional cases, private) datasets are nominated and designated as National Interest Datasets (NIDs).

This process should be public, driven by the National Data Custodian, and involve:

- The National Data Custodian accepting nominations for NIDs, assessing their public interest merits and, after consideration by the Government, referring selected nominations to a public scrutiny process. Designation would occur via a disallowable instrument on the recommendation of the National Data Custodian.
- The establishment of a parliamentary committee, or addition of such a role to the work of an existing parliamentary committee, to conduct public scrutiny of nominations for NIDs.

The process of nomination should be open to the States and Territories in order to cover linked datasets.

This process should be in place by the end of 2018, as part of the legislative package to implement these reforms.

RECOMMENDATION 7.2

In considering nominations for National Interest Datasets (NIDs), the National Data Custodian's public interest test should establish that through sharing or release, the designation of a dataset would be likely to generate significant *additional* community-wide net benefits beyond those obtained by the original data holder.

Once designated, NIDs that contain non-sensitive data should be made available for immediate release.

NIDs that include data on individuals would be available to trusted users only in a manner that reflects the accreditation processes of the relevant Accredited Release Authority, as established and updated by the National Data Custodian, to respect privacy and confidentiality.

Where data from the private and/or not-for-profit sectors is recommended to be included in a NID, the analysis prior to designation should specifically note the ways the designation addresses genuine commercial sensitivity associated with the information and costs (including those related to ongoing dataset maintenance).

RECOMMENDATION 7.3

Trusted users should be accredited by the relevant Accredited Release Authority (ARA) for access to those National Interest Datasets (NIDs) that are not publicly released, under processes accredited and updated as needed by the National Data Custodian.

Trusted users should be personnel from a range of potential entities that:

- have the necessary governance structures and processes in place to address the risks of inappropriate data use associated with particular datasets, including access to secure computing infrastructure, and
- have a signed legal undertaking that sets out safeguards for data use and recognises relevant privacy requirements.

The default position should be that after applicants and their institution establish capability to respect the processes and obligations of the ARA's accredited standard, an individual researcher from one of these organisations would be readily approved for access.

For trusted users of NIDs, this status should provide an ongoing access arrangement to specified unreleased datasets that would only cease on completion of a researcher's engagement with their relevant institution, or a loss of trust in the user or their organisation (via processes also established in accreditation of the ARA by the National Data Custodian).

RECOMMENDATION 7.4

The Australian Government should make provision, in select circumstances as approved by the funding Minister, for the National Data Custodian to pay for access or linkage to private sector datasets (Recommendation 9.4).

Equally, the National Data Custodian may consider applying charges for access to National Interest Datasets where this would not be inconsistent with the public interest purpose of the National Interest Dataset.

It is expected this would not be a common occurrence, in either case.

Legislative reform

RECOMMENDATION 8.1

New Commonwealth legislation — the Data Sharing and Release Act — should be passed drawing on the full range of Commonwealth powers to regulate digital data, in order to authorise the better sharing and release of data.

The new Act should also establish the Comprehensive Right of consumers to access their data from government and private data holders alike, for the purposes of improving the services that are offered to them by alternative providers.

RECOMMENDATION 8.2

The Data Sharing and Release Act should establish the risk-based approach to data sharing and release and accompanying institutional frameworks.

- All non-sensitive data held by agencies and Accredited Release Authorities (ARAs) should be explicitly presumed to be made public, consistent with the Australian Government's Public Data Policy Statement.
- Data custodians and ARAs would be authorised to provide sensitive data to trusted users in a secure environment, with de-identification where necessary for risk management of the data.
- The National Data Custodian should have the authority to issue guidance on how the risks of *all* sharing of identifiable data between entities should be managed. This guidance should be updated where it judges the risks have shifted.

RECOMMENDATION 8.3

The Data Sharing and Release Act (DSR Act) would, where possible, override secrecy provisions or restrictions on use that prevent original custodians actively providing access to data to other public sector data custodians and Accredited Release Authorities (ARAs).

Access should be governed by Data Sharing Agreements that embed the trusted user principles, actively assist data sharing and create clarity of understanding amongst the parties. The National Data Custodian (NDC) should issue a model Data Sharing Agreement early in its life, and update it from time to time.

The DSR Act should establish modern, clear and supportive standards — the new ‘rules of the game’ — for data sharing and release. The Commonwealth Privacy Act would continue to apply, as well as any residual obligations emanating from the original data custodian’s legislation.

Existing protections would remain on datasets that do not utilise the DSR Act, in order to ensure there is no gap between the accountability obligations on original public sector data custodians and the ARA.

In limited exceptional circumstances as the DSR Act transitions to becoming nationally effective, it may be necessary to provide access to data shared under the new Act to a party that has yet to adopt its provisions. The NDC should be provided with the power to use a disallowable instrument to allow access or sharing for such transitional purposes.

RECOMMENDATION 8.4

The Australian Government’s Protective Security Policy Framework (and equivalent State and Territory policies) should be amended to recognise that the risk and therefore the classification needed for data can be reduced by:

- transforming a dataset, for example through de-identification, such that the risks of misuse on dataset release are reduced
- only making the transformed data available to trusted researchers in a secure computing environment, with usage monitored and output checked for disclosiveness.

This would align the Protective Security Policy Framework with the current legal environment.

The Australian Government should consider doing this as part of its response to the Belcher Review.

RECOMMENDATION 8.5

Legislative reform to implement the Commission's recommendations would need to be undertaken in two parts, moving forward together:

- the first part is the passage of the Data Sharing and Release Act (DSR Act) itself, that authorises to the greatest extent practical in a single statute, the sharing and release of data for the purposes of the Act and removes existing Commonwealth and State restrictions on integrating, linking and research uses of datasets by Accredited Release Authorities
- the second part is a further legislative amendment process that may be necessary, depending on the particular characteristics of, for example, National Interest Datasets, in order to address residual restrictions on the use of specific datasets that were not able to be effected by the DSR Act itself.

The National Data Custodian should be asked to identify residual legislative restrictions that need removal in its consideration of National Interest Datasets.

RECOMMENDATION 8.6

The Data Sharing and Release Act (DSR Act) should have national reach — to create a simplified and transparent one-stop location for a national framework for data volunteered, declared or acquired for inclusion under the DSR Act.

The Act should allow for the acquisition of private datasets via disallowable instruments as part of the process of creating National Interest Datasets (NIDs). Acquisition should only occur on just terms after parliamentary scrutiny determines the benefits are demonstrable.

An initial set of NIDs should be identified by the National Data Custodian to accompany the DSR Bill, following processes to establish additionality and public interest.

The DSR Act should apply Commonwealth privacy legislation to datasets managed by Accredited Release Authorities where feasible. It should be drafted with reference to (and with the intention of being consistent with) the *Data Sharing (Government Sector) Act 2015* (NSW) and the *Public Sector (Data Sharing) Act 2016* (SA) to the extent possible.

RECOMMENDATION 8.7

The Australian Competition and Consumer Commission (ACCC) and the Office of the Australian Information Commissioner should enter into working arrangements with each other, industry ombudsmen and other relevant bodies at all levels of government to support a 'no wrong door' approach to how individuals (including small businesses) pursue complaints or queries regarding their rights as consumers to data held on them.

Where an industry data-specification agreement (Recommendation 5.2) seeks to use a recognised industry ombudsman to address consumer complaints, this should be considered by the ACCC as part of its acceptance or rejection of a proposed industry agreement.

Data transformation and pricing

FINDING 9.1

There is no single pricing approach that could act as a model for guiding public sector data release decisions.

The identification by agencies of the grounds for undertaking each release would have a direct bearing on the choice of price approach.

Cost recovery, long considered to be the default option in the public sector, is only one of a range of approaches and not necessarily to be preferred.

RECOMMENDATION 9.1

The emphasis for government agencies in handling data should be on making data available at a 'fit for release' standard in a timely manner. Beyond this, agencies should only transform data beyond the basic level if there is a clearly identified public interest purpose or legislative requirement for the agency to undertake additional transformation, or:

- the agency can perform the transformation more efficiently than either any private sector entities or end users of the data; and
- users have a demonstrable willingness to pay for the value added product; and
- the agency has the capability and capacity in-house or under existing contract; and
- the information technology upgrade risk is assessed and found to be small.

RECOMMENDATION 9.2

The pricing of public sector datasets for public interest research purposes should be the subject of an independent review.

RECOMMENDATION 9.3

Minimally processed public sector datasets should be made freely available or priced at marginal cost of release.

Where data has been transformed, the transformed dataset may be priced above the marginal cost of release. Data custodians should experiment with low prices initially to gauge the price sensitivity of demand, with a view to sustaining lower prices if demand proves to be reasonably price sensitive.

RECOMMENDATION 9.4

Funding should be provided to agencies for the curation and release of those datasets determined through the central data agencies' public request process (Recommendation 6.5) to be of high value with a strong public interest case for their release. This funding should be limited and supplemental in nature, payable only in the event that agencies make the datasets available through public release.

Funding would also be required for the Office of the National Data Custodian, for functions undertaken by Accredited Release Authorities and, in some cases, for the purchase and ongoing maintenance of National Interest Datasets. Additional responsibilities required of the Australian Competition and Consumer Commission in regard to the Comprehensive Right should also be resourced.

Aside from these purposes, no additional supplementary funding appears warranted for agencies' activities related to their data holdings as a consequence of this report.

Implementation of the new Framework

RECOMMENDATION 10.1

The Australian Government should engage actively with the community on matters related to data availability and use.

At a minimum, the National Data Custodian should regularly convene forums for consultation, to ensure community concerns about increased use of data are addressed.

RECOMMENDATION 10.2

The Australian Government should set an ambitious — but realistic — timeline for implementation of the Commission’s recommended reforms.

A set of actions in this Report can be completed in 2017, to ensure they deliver benefits to the community in the short term.

Passage of the Data Sharing and Release Act and supporting Part 2 amendments for an initial suite of National Interest Datasets should be in place by the end of 2018.

A central agency with data responsibility should actively support the progress made against the implementation plan until the Office of the National Data Custodian is legislatively established.

Once established, the National Data Custodian should assume responsibility for monitoring and evaluating the effects of the new data Framework, reporting annually on progress and with a formal evaluation after three years’ experience of the Framework’s reforms.

RECOMMENDATION 10.3

Government agencies should adopt and implement data management standards to support increased data availability and use as part of their implementation of the Australian Government’s Public Data Policy Statement.

These standards should:

- be published on agency websites
- be adopted in consultation with data users and draw on existing standards where feasible
- deal effectively with sector-specific differences in data collection and use
- support the sharing of data across Australian governments and agencies
- enable all digitally collected data and metadata to be available in commonly used machine-readable formats (that are relevant to the function or field in which the data was collected or would likely be most commonly used), including where relevant and authorised, for machine-to-machine interaction.

Policy documents outlining the standards and how they would be implemented should be available in draft form for consultation by the end of 2017, with standards implemented by the end of 2020.

Agencies that do not adopt agreed sector-specific standards would be noted as not fully implementing the Australian Government’s Public Data Policy and would be required to work under a nominated Accredited Release Authority to improve the quality of their data holdings.

RECOMMENDATION 10.4

The private sector is likely to be best placed to determine sector-specific standards for data sharing between firms, where required by reforms recommended under the new data Framework.

In the event that cooperative approaches to determining standards and data quality do not emerge or adequately enable data access and transfer (including where sought by consumers), governments should facilitate this.

Key points from chapters

Chapter 1 Australia's data landscape

Key points

- Data refers simply to a collection of material, which can include characters, text, words, numbers, pictures, sound or video. Data may be stored digitally or in hard copy formats, with digitisation enabling data to be copied, stored and transferred rapidly.
- New sources of data — as varied as social media sites, smart mobile devices and sensors fitted to physical objects (the 'Internet of Things') — continue to emerge and expand. Digital data, a source of considerable potential value, is being collected ubiquitously.
- The extraordinary capabilities of data analytics and the increasing ability to link previously separate datasets are compounding the usefulness of new data sources, offering important opportunities for better-informed decision making by individuals, businesses and governments, and for research breakthroughs.
- The frameworks and protections for data collection and access, developed prior to sweeping digitisation, require serious re-examination. As one example only, privacy law is neither the only lens, nor even the best, through which to view the use of an asset such as data.
- A shift to viewing data as an opportunity, not necessarily a threat, is a global phenomenon.
- There can be many different competing interests in a particular dataset, including: the subject of the data (such as an individual, who is often also the source); the parties who collect, aggregate and analyse the data; and those who commission these actions. Clarity about these interests is essential to allow Australia to harness the full value of its data.
 - The line between what is 'personal' data and what is not is blurred both legally and practically, as shown by a recent Federal Court ruling. The readiness of individuals to share information about themselves on social media, and other avenues such as loyalty programs, may indicate that social appetite for some types of data use are changing.
- A common misperception is that privacy laws — or, indeed, the privacy policies of individual organisations — give individuals ownership over data created by or about them.
 - In Australia, no one 'owns' data, although copyright law may apply in limited circumstances. Privacy legislation, the primary generic tool offering individuals some control, regulates how personal information is collected, used and disclosed.
 - In a world increasingly making use of the data of *others*, the primary unaddressed question should be: for how long will the public — the source of most of this information — trust a structure in which their actual rights are mainly limited to privacy?
- An enormous range of information is collected by governments, researchers and businesses about individuals and their activities, institutional and economic structures, and the built and natural environments. However, there is less publication — or controlled sharing — of this information than would help achieve widespread benefits for the community.
- This Report offers guidance on how governments may generate community acceptance of the processes, costs and risks associated with enhanced data use, and to do so where benefits may be most evident.

Chapter 2 Opportunities enabled by data

Key points

- In less than two decades, extraordinary developments in computing power, data generation and algorithms that can detect patterns and preferences previously indiscernible, have enabled new business models and opportunities. Individuals, businesses, governments and the broader community have all benefited from these changes.
- Numerous examples of innovative uses of data are offered by this Report and submissions to it. Many such innovations would be unimaginable without the ability to collect and interpret large volumes of data.
- Private sector data owners are leading the way in finding innovative uses for data.
- Governments across Australia also hold lots of data, but are typically not using it to the extent that opportunities being taken overseas exemplify, and lack a comprehensive plan to do so in most cases. That said, there are some notable examples of governments seeking to use data in innovative, value-adding ways, most of which rely on the personal leadership of a handful of individuals rather than systemic commitment.
- Opportunities to make more productive use of datasets — such as by applying analytical algorithms or linking them to other datasets — and the benefits that could be achieved, are largely unknown until the data sources themselves are made known and a wide range of users have had opportunities to investigate the data. This underscores the substantive argument in favour of greater data access.
- The distribution of benefits from greater access to data would be widespread, but the Commission is of the view that it is individual consumers who have the most to gain, collectively, from action on Australia's data sharing and release arrangements.

Chapter 3 What holds us back?

Key points

- A key to achieving the many potential benefits of data use will be building and retaining community trust in how data is managed and used and building a shared understanding of the benefits that flow from better data access and use, including by consumers themselves.
- Community surveys indicate some concerns about the privacy and security of personal information. Such concerns are compounded by individuals' lack of knowledge of what data about them is collected, how it is used and what rights they have over such data, and the often impenetrable nature of data-related terms and conditions.
- Legislation restricting access to data was formulated up to a century ago, and much is no longer fit for purpose. The primary legal impediment to more effective use of data is typically *not* the Privacy Act, but regulations and guidelines specific to the field in which the data is collected.
- A culture of risk aversion among public servants has led to overly cautious interpretation of relevant legislation, lack of willingness to make it known that some data exists (lest the agency be asked for it), and complex and lengthy approval processes for data access (including duplicative ethics committee requirements).
- A lack of national leadership has contributed to piecemeal bureaucratic processes for data sharing and release. Sharing of restricted access data between government agencies and certain trusted parties has been limited at best. There are numerous examples of data siloing, despite the obvious benefits of data sharing.
- The extent of productive linking and integrating of datasets varies substantially across jurisdictions, but is generally inadequate when viewed against the potential opportunities or practices in some other countries.
- For the vast majority of publicly funded research, neither the public nor most of the research community has access to datasets generated by projects, despite there being a clear public interest in this occurring. Data releases are infrequent and very dated. Destruction of datasets — a particular policy of the Commonwealth — increases the costs of data re-use.
- Technical challenges are used to justify risk averse practices: the risk of data breaches and re-identification of de-identified personal data; fragmented data collection and release; lack of common standards; and a shortage of skills and dedicated resources. None of these are driven by increased data sharing, but are nevertheless cited as reasons not to do so.
- The collection and use of data allows businesses to improve their competitiveness, but on their terms. While individuals can currently see their own data, there is no obligation on any data collector to provide this in a useful form.
- Better access to, and the opportunity for individuals to use data held about them, would likely spark additional competition and innovation to the benefit of consumers.

Chapter 4 A way forward: what we must aim at

Key points

- Data is a strategic asset with great potential and should be treated and managed as such.
- Increased data collection and use is an economy-wide tidal wave already upon us. While it moves inexorably forward, we must move policy and practice forward in a similarly comprehensive fashion.
 - Piecemeal reforms would not achieve the same benefits, nor would they enable effective management of data risks.
- The Commission's recommended reforms are designed to create a scalable, risk-based framework for data sharing and release that will support Australia well into the future.
- Key elements of the reforms include:
 - a legislative framework designed to provide a clear and modernised approach to data access and use
 - new rights for consumers to enable them to share in the benefits of data — these demand-focused reforms would drive better choices and more competitive markets
 - establishing a scalable, risk-based institutional framework to allow integration and sharing or release of Australia's data
 - recognising that some datasets are of such significance they should be treated as national assets.
- The reforms are designed to work together to provide mutual assurance and benefit. Splitting off one principal element from another risks a lack of community trust and support if opportunities for individuals are abandoned; and a serious loss to community welfare if public dataset integration, sharing and release is delayed.
- Building and maintaining a social licence is front and centre of our reform package. Social licence will develop if people:
 - have a sound basis for believing in the integrity and accountability of institutions (public and private) managing data
 - have an inalienable ability to *choose* to participate in extracting benefit from data sharing
 - have some control over how their data is used
 - better understand the potential community-wide benefits of data sharing and use.
- Robust institutional and governance arrangements regarding data use build public trust and maintain incentives for better practice over time. Well-resourced and capable institutions are critical.

Chapter 5 **New competition policy — a right to use your data**

Key points

- Maintaining a social licence for wider data use, both public and private, can be actively supported by offering individuals the opportunity to participate, as firms and governments do, in accessing and re-using their own data.
- Rights to use data will give better outcomes for consumers than ownership: the concept of *your* data always being *your* data suggests a more inalienable right than one of ownership (which can be contracted away or sold). And in any event, consumers do not own their data in Australia.
- A new Comprehensive Right for consumers (including small/medium-sized businesses) would provide greater insight and control for individuals over how data that is collected on them is used. Consumer data for this Right should be defined broadly and with a focus on desired outcomes, but with an opportunity for inclusion of data that is merely imputed to be about the consumer.
- The Comprehensive Right is significant in a policy sense beyond its ability to support a social licence for better data use economy-wide. It may offer the capacity to underpin a new wave of competition policy, similar in its catalytic effect to the Hilmer reforms of the 1990s.
- Under the new Right, all consumers would have a right to obtain a machine-readable copy of their own digital data, provided to them and/or to a nominated third party, such as a potential new service provider.
 - Existing privacy provisions to view and request edits or corrections to personal information would remain, with the new Right also applying these to consumer data.
 - Consumer data would be a joint asset between the individual consumer and the entity holding the data. Exercise of the Right by a consumer would not alter the ability of the initial data holder to retain and keep using the data.
- Determining what data the new Right should apply to will be controversial at times. The coverage of the Right and transfer method should be agreed within each industry through a standard-setting and data-specification process, the outcome of which would require the Australian Competition and Consumer Commission's approval.
 - Data available for transfer must, at a minimum, be sufficient to enable consumers to meaningfully transfer their custom and obtain service from another supplier.
 - Absence of industry agreement would mean the consumer data defaults to a broad definition.
- Participation in comprehensive credit reporting has been low to date and the associated benefits are far from fully realised. A target for participation of 40% of all active credit accounts provided by Australian Securities and Investments Commission-licensed credit providers should be set for 30 June 2017. Legislation to mandate participation should be circulated for consultation by the end of 2017, if the target is not met.

Chapter 6 Sharing and releasing data for community benefits

Key points

- A structure for data sharing and release should be introduced that would allow access arrangements to be ‘dialled’ up or down according to the nature of risks associated with different types of data, uses and use environments.
- The structure would cover a wide range of data, including identifiable data that can be shared with trusted users, de-identified data, and open access data.
- The recommended changes to the current handling of the public sector’s data assets are designed to maintain public confidence and build a social licence for data re-use.
- Implementation and ongoing monitoring would be vested with a new statutory office holder, the National Data Custodian (NDC).
 - The NDC would administer the Data Sharing and Release Act, recommend valuable datasets for designation as National Interest Datasets, provide technical guidance and direction to the data system, accredit and coordinate Accredited Release Authorities (ARAs), report on data use ‘good news’ and breaches, and manage broader ethical considerations within the data release and sharing system.
 - The NDC would not, however, hold data itself or decide on its release or re-use.
- An integral part of the Commission’s recommended data reform Framework is the introduction of a network of ARAs.
 - ARAs would be hubs of sectoral expertise in data curation, de-identification and linkage, would implement a risk-based approach to the broader sharing and release of data through formal, contemporary, NDC-reviewed risk management procedures.
 - ARAs would generally be nationally focused and may be located within the Commonwealth, State or Territory public sectors or be other publicly funded entities that have the necessary expertise, focus and governance structures.
 - ARAs would be funded to undertake data curation, management, linkage, storage and release, with funding determined by the NDC. ARAs could also receive revenue from project work undertaken for clients.
- There is a compelling public interest to see greater use made of public-funded researchers’ unique datasets. Progress by individual research institutions that receive government funding in making their research data and metadata widely available to others should be openly published by those institutions, with reference to past performance.
- Existing exceptions in privacy legislation allowing the sharing of personal information for health and medical research purposes without consent should be extended to other areas of public interest research.
- The requirement to destroy linked datasets and statistical linkage keys at the completion of data integration projects should be abolished. It is akin to book burning.

Chapter 7 Getting value from Australia's national interest datasets

Key points

- Wider release of data and more effective sharing by governments would likely trigger significant investment (private as well as public) and improvements in national welfare. However, determining which datasets could lead to such improvements remains a serious practical issue for governments.
- Existing government data initiatives, such as data.gov.au, should be leveraged as part of the broadest sharing and release efforts. A framework is needed to formalise how such efforts would be implemented and managed, and is provided in this Report.
- Datasets of high value have a number of distinct characteristics, including that they are unique (or cannot be readily replicated), are of high quality, have a high degree of coverage in the relevant population, and are up-to-date or updated regularly.
- National interests additionally require that coverage is oriented towards nationally significant subject matter. Data with these combined characteristics are likely to generate spillover benefits for the community and should be designated as National Interest Datasets (NIDs).
 - Prioritising selection of NIDs would require flexibility and discretion. A parliamentary committee would be a suitable vehicle to scrutinise this process, rather than try to define national interest in legislation. Engagement with democratic representation would assist in maintaining social licence.
 - The process for selecting datasets for designation as NIDs would be open to the public. As the funder of NIDs, the Australian Government would be the arbiter of their selection.
 - Initial candidates for consideration as NIDs include existing public sector datasets that provide registers of businesses, services or assets, or record activity in key areas of public expenditure, such as health and education, as well as datasets held in private entities that are regulated in the public interest and/or receive public funding.
 - It is expected that other datasets would be nominated as NIDs by a range of parties, including State and Territory governments (preferably acting collectively), private sector entities and not-for-profit organisations. Incentives to do so include provision by the Commonwealth of ongoing funding for integration and maintenance, a desire for national linkage and access, and the removal of current restrictions on sharing and release, replaced by fit-for-purpose safeguards.
 - Access to NIDs would be provided via the processes outlined in chapter 6, by a designated Accredited Release Authority (ARA).
- The National Data Custodian (NDC) would lead the process to identify and develop the case in support of NIDs; and ensure the chosen ARA makes NIDs widely available.
 - The processes used by the NDC should be transparent and risk-based, and a relationship not unlike that of the Auditor General with the Public Accounts Committee is anticipated.
 - A focus on the net benefits likely to accrue from designating a dataset as NID, the possible impacts on intellectual property, impacts on incentives to continue collection, and a detailed consideration of the costs to the collecting party or parties, should be points of primary focus in deliberation.

Chapter 8 A modernised regulatory framework

Key points

- Legislation, in the form of the Data Sharing and Release Act, should be introduced to create consistent rules for improved data sharing and release.
- The Data Sharing and Release Act serves two major purposes:
 - First, as a clear and unambiguous signal of the shift in approach of the Commonwealth Government and parliament to the data issues uncovered in this report, a crucial step to achieving cultural change amongst a myriad of data custodians.
 - Second, as a structural framework, principles-based and outcomes-focused as far as practical, to give authority and guidance to effective and responsible use of information in a data-rich future.
- The Data Sharing and Release Act would require a risk-based approach to improved data sharing and release.
- It would, for the purposes of other legislation that impedes consideration of improved sharing and release, authorise the sharing of data within the public sector and with ARAs.
- New institutional arrangements to promote trust and confidence in the data sharing and release system are centred on improved capabilities, accreditation procedures for both users and custodians, and most particularly the National Data Custodian.
- National Interest Datasets would overcome impediments to the effective integration, sharing and use of data of national significance, currently hindered by multiple legislative barriers.
- Privacy is an important human right and existing protections are retained. The Data Sharing and Release Act should, where it deals with release of personal information (as defined in privacy legislation), operate subject to existing Commonwealth privacy legislation.
- To ensure that individuals (including small businesses) can participate in data sharing opportunities, a new definition of data — consumer data — would be created to cover all digital information to deliver the ability of consumers to utilise their data, as outlined in chapter 5. The changes in consumer data rights should be supported proactively by existing consumer-oriented regulators.
- The Privacy Act is not an effective vehicle for the reforms covered in this report, particularly when it comes to consumer data. Privacy regulators seem unconvinced of the need, and emphasised to us their preference for the human right of privacy over consumer interests. Whereas from a national welfare perspective, as required in the Productivity Commission's Act, it is evident that privacy is one aspect of data use, not pre-eminent other than that it is incumbent.
- The Data Sharing and Release Act should, to the maximum extent available under Commonwealth powers, establish a national framework, and offer cooperating jurisdictions (and occasionally private datasets) the ability to become integrated and accessible for research and other authorised purposes.
- The benefits of a consistent framework, that is scalable and adaptive over time, have generally been widely recognised by stakeholders responding to the Draft Report.

Chapter 9 Transformation and pricing decisions

Key points

- Opportunities for harvesting data and creating new business ventures are important determinants of the extent to which organisations choose to transform data, and/or share it with other parties. The transformation of data to enhance its value, and its sharing with others, has resulted in major structural change in some markets and businesses over the last decade. Structural change in government has been much less substantial.
- Public and research sector datasets also offer attractive opportunities for new and innovative private sector services, but government agencies and research bodies have limited ability to determine the value of data they might release.
- Assessing value could be important for assessing the case for additional dataset processing, such as tailoring datasets for specific uses.
- Processing of data that adds value and takes a dataset beyond the standard required internally and for other regulatory requirements, should only be undertaken by government agencies when:
 - there is a previously-unaddressed public interest purpose clearly identified by the agency, and accepted by the government, for the agency to undertake additional transformation and make the value-added data available; or
 - the agency can perform the transformation more efficiently than either users of the data or private sector intermediaries; and potential users of the data have a demonstrable willingness to pay; and agencies have the capability in-house or under contract with a third party; and the information technology upgrade risk is assessed and found to be small.
- Beyond this, government agencies should refrain from additional transformation of datasets. The delay incurred by agencies in doing so prior to release (or sharing) can be substantial, and data users generally have a preference to access data ‘as is’ and in a timely manner.
- There are various approaches for pricing public sector data, ranging from free provision and marginal cost pricing to commercial pricing. Which approach is most suitable will vary according to user demand and agency capability to act commercially.
 - For a given level of data quality, making data freely available in a timely way will maximise use and hence deliver the highest level of social benefits. But it will increase the net cost to government of data release.
 - Where agencies undertake substantial transformation because it meets the principles above, there are strong grounds for passing these additional costs on to data users.
 - An exception may be where data is used for research in the public interest. Pricing of data for the publicly-funded research community should be the subject of a separate review.
- Maintaining and increasing the availability of public sector data is not costless. Sharing or release of minimally processed datasets should be funded by agencies from existing budgets. For datasets that have significant and demonstrable public interest and are to be made more available, additional funds should be provided to agencies to ensure the quality of dataset curation.

Chapter 10 Implementing the new data Framework

Key points

- Implementation processes for Australia's new data Framework must be as open and transparent as possible. All beneficiaries — the community, governments and businesses — should be involved in an exchange of views around data handling and use, in a manner that raises confidence in Australia's data ecosystem. Changes of this order cannot be suddenly launched into the public arena.
- State and Territory Governments will have a major role to play in the new data Framework and must continue to be engaged actively.
 - Some have already made substantial progress, by introducing data sharing legislation and creating agencies that specialise in data. The Commission's recommended reforms build on this progress to promote data sharing within and between jurisdictions.
- The Australian Government should set an ambitious timeline for reform implementation. The Commission envisages that some of its recommendations can — and should — be implemented immediately, including:
 - abolishing the requirements to destroy linked datasets that include Commonwealth data
 - accrediting state-based linkage units to link Commonwealth data
 - creating registers of data held by publicly funded entities.
- Before the end of 2017, the Australian Government should establish the Office of the National Data Custodian (NDC) via administrative means.
 - The NDC can add substance to the debate proposed above, and become a focal point for interaction with parties essential to the passage of the Data Sharing and Release Act.
 - It can also commence developing the processes for accrediting ARAs and trusted users, and enabling stakeholders to debate nominations for National Interest Datasets, to be implemented on passage of the legislation.
- At the same time, the Australian Government should lift the profile of the national conversation on data, using the proposals and evidence in this Report as a base. The emphasis should be on all stakeholders gaining greater certainty from the data reform process.
- The Australian Government should work towards passing the Data Sharing and Release Act before the end of 2018.
- A central agency with data responsibility should be tasked with monitoring the reform implementation process and remain in that role until the legislative steps are completed. It should report publicly on progress at the end of 2017 and again in 2018, to set the tone for maintaining this process as a transparent one. From 2019 onwards, the NDC (presuming passage of legislation) should take on this role.

References

- ABS (Australian Bureau of Statistics) 2016e, *Personal Fraud, Australia, 2014–15*, Cat. 4582.0, Canberra.
- ACMA (Australian Communications and Media Authority) 2012, *Location services, personal information and identity: Exploratory community research*, Australian Government, Canberra, <http://www.acma.gov.au/theACMA/Library/researchacma/Research-reports/here-there-and-everywhere-consumer-behaviour-and-location-services> (accessed 4 July 2016).
- 2013a, *Privacy and personal data*, Occasional Paper #4, Emerging issues in media and communications, Australian Government, Canberra, <http://www.acma.gov.au/theACMA/About/The-ACMA-story/Connected-regulation/privacy-and-digital-data-emerging-issues> (accessed 4 July 2016).
- CBO (Congressional Budget Office (US)) 2008, *Evidence on the Costs and Benefits of Health Information Technology*, United States Congress, Washington D.C.
- Chaudhry, B., Wang, J., Wu, S. and Maglione, M. 2006, ‘Systematic review: impact of health information technology on quality, efficiency, and costs of medical care’, *Annals of Internal Medicine*, vol. 144, no. 10, pp. 742–752.
- Directivity, Citrus and First Point Research 2015, *For love or money: 2015 consumer study into Australian loyalty programs*, Melbourne, <http://www.theloyaltypoint.com.au/for-love-or-money-2015> (accessed 25 July 2016).
- EMC Corporation 2014, *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things*, <http://www.emc.com/leadership/digital-universe/2014iview/index.htm> (accessed 14 July 2016).
- IBM 2016, *American Sleep Apnea Association and IBM Launch Patient-led Sleep Study App; First ResearchKit App on Watson Health Cloud*, 2 March, <https://www-03.ibm.com/press/us/en/pressrelease/49275.wss> (accessed 27 April 2016).
- OAIC (Office of the Australian Information Commissioner) 2013a, *Community Attitudes to Privacy Report: 2013*, Australian Government, Sydney.
- OECD (Organisation for Economic Cooperation and Development) 2015b, *OECD Reviews of Health Care Quality: Australia 2015*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264233836-en> (accessed 18 May 2016).
- Research Australia 2016, *Australia Speaks! Research Australia Opinion Polling 2016*, <http://researchaustralia.org/reports/public-opinion-polling/> (accessed 20 September 2016).

Srinivasan, U., Rao, S., Ramachandran, D. and Jonas, D. 2016, *Flying blind - Australian Consumers and Digital Health*, Australian Health Data Series: Volume 1, Capital Markets Cooperative Research Centre (CMCRC).

UN ECE (United Nations Economic Commission for Europe) 2015, *Big Data*, United Nations, Geneva, <http://www1.unece.org/stat/platform/display/msis/Big+Data> (accessed 1 August 2016).