DATA LINKAGE EXPERT ADVISORY GROUP

Developing a whole-of-Government data linkage model

A review of Western Australia’s data linkage capabilities

December 2016
Expert Advisory Group

Chair
Professor Peter Klinken
Chief Scientist of Western Australia

Members
Professor Fiona Stanley AC
Patron and Founding Director, Telethon Kids Institute
Distinguished Research Professor, The University of Western Australia

Mr Giles Nunis
Western Australian Government Chief Information Officer

Assisting Members
Professor Bruce Armstrong
Emeritus Professor, School of Public Health
University of Sydney

Professor Louisa Jorm
Foundation Director, Centre for Big Data Research in Health
University of New South Wales

Project Team

Project Director
Ms Fiona Roche
Executive Director, Office of Science
Department of the Premier and Cabinet

Project Manager
Ms Kathryn James
Director, Office of Science
Department of the Premier and Cabinet

Project Officer
Ms Karina Schaap
Senior Policy Officer, Office of Science
Department of the Premier and Cabinet

Project Researcher
Mr Adam Byass
Graduate Officer, Office of Science
Department of the Premier and Cabinet
[INTENTIONALLY LEFT BLANK]
Table of Contents

Glossary ........................................................................................................................................... 7
Executive Summary ......................................................................................................................... 14
Recommendations ............................................................................................................................ 18
PART 1 – INTRODUCTION AND BACKGROUND .............................................................................. 22
Chapter 1 - Introduction .................................................................................................................. 23
Chapter 2 – Background .................................................................................................................. 27
  2.1. Definition of data linkage ...................................................................................................... 27
  2.2. Overview of the data linkage process in WA .......................................................................... 28
  2.3. History of data linkage in WA .............................................................................................. 29
  2.4. National data linkage landscape and initiatives .................................................................... 34
  2.5. Benefits of data linkage ........................................................................................................ 37
  2.6. International examples ......................................................................................................... 41
PART 2 – KEY ISSUES & RECOMMENDATIONS .............................................................................. 45
Chapter 3 – Governance .................................................................................................................. 46
  3.1. Current governance arrangements ....................................................................................... 47
  3.2. Inclusiveness and consultation .............................................................................................. 47
  3.3. Transparency ....................................................................................................................... 48
  3.4. Future governance – a centralised approach ........................................................................ 49
Chapter 4 – Privacy .......................................................................................................................... 53
  4.1. Privacy of personal information and data linkage ............................................................... 54
  4.2. Current privacy measures for data linkage .......................................................................... 54
  4.3. Future privacy measures for data linkage ............................................................................ 55
Chapter 5 – Ethics ............................................................................................................................ 59
  5.1. Current ethics review processes ........................................................................................... 60
  5.2. Future ethics review processes ............................................................................................. 63
Chapter 6 – Data Availability and Access ....................................................................................... 65
  6.1. Data access and data linkage ............................................................................................... 66
  6.2. Access to data held by WA Government agencies ................................................................. 66
  6.3. Access to data collected by the Commonwealth Government ........................................... 71
  6.4. Access to data held external to Government ...................................................................... 72
Glossary

**ABS** – Australian Bureau of Statistics

**AIFS** – Australian Institute of Family Studies

**AIHW** – Australian Institute of Health and Welfare

**ARC** – Australian Research Council

**ATO** – Australian Tax Office

**BC** – British Columbia

**CARES** – Custodian Administered Research Extract Server. A Data Linkage Branch initiative that streamlines linked data extraction, quality control and delivery services.

**CDL (see PHRN)** – Centre for Data Linkage

**CHeReL** – The NSW Centre for Health Record Linkage. Data linkage service connecting data from the health, education, human services, justice and transport sectors aimed at enabling research and evaluation. The service was established in 2006 as a dedicated data linkage unit and is managed by the NSW Ministry of Health.

**COAG** – Council of Australian Governments

**CVDL** – The Centre for Victorian Data Linkage

**DAC** – The NSW Data Analytics Centre. DAC facilitates data sharing between agencies to inform whole-of-Government evidence-based decision making. It does this by leveraging internal and external partnerships so that the right capabilities, tools and technologies are applied. Established under the *Data Sharing (Government Sector) Act 2015 (NSW)*.

**DLAB** – Data Linkage Advisory Board. Governance body responsible for: a) strategic growth of the WA data linkage infrastructure, b) privacy considerations of linked health data, c)
budget funding and review, d) IT infrastructure, and e) access to linked data and training needs. The DLAB ceased in 2012.

**DLB** – Data Linkage Branch. The team at the Department of Health who are responsible for developing and maintaining the WA Data Linkage System, performing data linkage, and the facilitation of access to linked data. Data can be requested for ethically approved research, planning and evaluation projects which aim to improve the health of Western Australians.

**DoH** – Department of Health

**DPC** – Department of the Premier and Cabinet

**DPMC** – Department of the Prime Minister and Cabinet

**DPP** – Developmental Pathways Project. Collection of research projects aimed at providing new knowledge to inform whole-of-Government intervention and prevention strategies designed to improve health and wellbeing, and developmental and social outcomes. The DPP was established by the Telethon Kids Institute in collaboration with The University of Western Australia (Crime Research Centre), Treasury, the Disability Services Commission and the Departments of Health, Education and Training, Child Protection, Community Development, Public Housing, Police and Justice.

**Data custodian** – The person within an organisation or agency formally assigned to collect, manage, secure and disclose a data set on a day-to-day basis at the direction of the data steward.

**Data extraction** – The act or process of retrieving data out of (usually unstructured or poorly structured) data sources for further data processing or data storage. The term data extraction is also applied when data is first imported into a computer from primary sources, such as measuring or recording devices.

**Data linkage** – A complex technique for connecting data records within and between data sets using demographic data (e.g. name, date of birth, address, sex, medical record number). Also known as ‘Record Linkage’ or ‘Linkage’.

**Data management** – The development and execution of architectures, policies, practices and procedures that properly manage the full data lifecycle needs of an organisation.
Data privacy – Refers to the challenge of utilising data while protecting individual’s privacy preferences and their personally identifiable information. Robust privacy protection is critical to ensuring the success and survival of any data linkage system.

Data quality (high and low quality data) – Refers to the level of quality of data. There are many definitions of data quality but data are generally considered high quality if they can be considered fit for their intended uses in operations, linkage, decision making and planning.

Data set – A collection of records containing similar items of information (data) about a number of record subjects; for example, a WA Births data set might contain many thousands of records, each containing the same pieces of information, such as name, place, and date of birth for infants born in WA over a particular period of time.

Data steward – The person within an organisation or agency formally assigned to set the strategic purpose, operation and disclosure model of a data collection.

De-identified – Related to the identifiability of records in a data set; where the identity of a person or organisation has been removed from the records and therefore is not immediately obvious nor can be reasonably ascertained using other sources of information. The National Health and Medical Research Council discourages the use of this term as its meaning can be unclear, however, this term is still used by a number of other authorities. See related term ‘unidentifiable’.

Director General – The most senior public servant in a Government Department.

EOI – Expression of interest

GCIO – Government Chief Information Officer

HREC – Human Research Ethics Committee. A group constituted and operating according to National Health and Medical Research Council guidelines that oversees the ethical practice of research within an organisation. The Department of Health’s HREC oversees principally the use of Department of Health personal health information collections in research.

**Individually identifiable data** – Data from which the identity of a specific individual can reasonably be ascertained (also commonly referred to as identifiable data).

KEMH – King Edward Memorial Hospital

LPNO – Linkage Project Record Number - DLB’s in-house record identifier, which is a unique character string assigned to every record loaded into the WA Data Linkage System.

Landgate – The Western Australian Land Information Authority is a statutory authority that operates under the business name of Landgate.

Linkage key – the data linkage process generates a set of indices sometimes called linkage keys that are stored by the Data Linkage Branch in a links table. The linkage keys are held separately from any personal demographic information and they enable related health records to be joined together for approved research projects.

MADIP – Multi Agency Data Integration Project. A Commonwealth Government cross-portfolio partnership aimed at improving accessibility to, and maximising the use of, public data.

MBS – Medicare Benefits Schedule. List of the medical services for which the Australian Government will pay a Medicare rebate, to provide patients with financial assistance towards the costs of their medical services.

MCHP – Manitoba Centre for Health Policy

Metadata – Defined as data, or information, that provides information about other data. Three distinct types of metadata exist: structural metadata, descriptive metadata, and administrative metadata.
NCRIS – National Collaborative Research Infrastructure Strategy. Part of the Commonwealth Government’s National Innovation and Science Agenda. $1.5 billion over ten years is currently committed to this program, which aims to spur development of new research systems, infrastructure and innovation across Australia.

NDIS – National Disability Insurance Scheme

NHMRC – National Health and Medical Research Council


NSW – New South Wales

NT – Northern Territory

NZ – New Zealand

Non-identifiable data – Data not labelled with individual identifiers or from which identifiers have been permanently removed (i.e. de-identified) and by means of which no specific individual can be identified. A subset of non-identifiable data are those that can be linked with other data so it can be known they are about the same data subject, although the person’s identity remains unknown.

OGCIO – Office of the Government Chief Information Officer

PBS – The Pharmaceutical Benefits Scheme. Program of the Commonwealth Government that provides subsidised prescription drugs to residents of Australia, as well as certain foreign visitors covered by a Reciprocal Health Care Agreement.

PHRN – The Population Health Research Network. A collaborative program, funded by National Collaborative Research Infrastructure Strategy and managed by The University of Western Australia, that involves Government agencies, universities and medical research institutes and is aimed at creating a national data linkage network that would enable existing health data from around the nation to be collected and made available for health related research purposes.
**Personal data** – Data which relate to an individual who can be identified – (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

**QLD** – Queensland

**Re-identifiable data** – Data from which identifiers have been removed and replaced by a code that can be used to link records in the data back to the identities of the record subjects.

**SA** – South Australia

**SAIL** – Secure Anonymised Information Linkage. SAIL is a Wales-wide research resource focused on improving health, well-being and services. SAIL receives core funding from the Welsh Government’s Health and Care Research Wales. A range of anonymised, person-based data sets are held in SAIL, and, subject to safeguards and approvals, these can be anonymously linked together to address important research questions.

**SHIP** – ScottisH Informatics Programme. Scotland-wide collaborative research programme bringing together the Universities of Dundee, Edinburgh, Glasgow and St Andrews with National Services Scotland to develop SHIP Principles for the collation, management, dissemination and analysis of electronic patient records.

**SLIP** – Shared Land Information Platform. An open data platform for location-based information coordinated by Landgate through the Western Australian Data Linkage System partnership. Thousands of data sets can be accessed through SLIP via a wide range of maps.

**SPUR** – Location technology hub powered by Landgate. SPUR provides simplified access to data, resources and contacts within Landgate’s system.

**SUFEX** – Secure Unified File Exchange. Secure file transfer service for the Population Health Research Network (PRHN) and its stakeholders. SUFEX is one of the options that data custodians can use to send files to the PHRN data linkage units and researchers.
**SURE** – Secure Unified Research Environment. An Australian remote-access computing environment that allows data users to securely access and analyse linked data files. Developed by the Sax Institute as part of the Population Health Research Network.

**Statistics Canada** – Main access point for linked social science and health survey data for Canadians.

**Statutory Body** – Also referred to as a Statutory Authority. An organisation that has been created by parliament.

**TAS** – Tasmania

**TDLU** – Tasmanian Data Linkage Unit

**TKI** – Telethon Kids Institute. One of the largest medical research institutes in Australia, comprising more than 500 staff and students, with a focus in multidisciplinary research.

**Treasury** – Western Australia’s Department of Treasury

**Unidentifiable** – Related to the identification of data sets/data items; where the identities of the persons/organisations that are the subjects of the data sets/data items are not immediately obvious and it is not plausibly possible to identify or re-identify any subject/person/organisation using other sources of information.

**UWA** – The University of Western Australia

**VIC** – Victoria

**WA** – Western Australia

**WADLS** – Western Australian Data Linkage System. This system is used to connect available health and other related information for the Western Australian population. This incorporates database tables holding demographic data and linkage keys, and the bespoke tools used to process, create, store and retrieve them.
Executive Summary

On 17 March 2016 this report was commissioned by the then Director General of the Department of the Premier and Cabinet (DPC) to review Western Australia’s (WA’s) data linkage capabilities, to maximise the State’s competitive advantages in health data linkage and to develop a whole-of-Government model that builds on these strengths for the future. The review was undertaken by the Data Linkage Expert Advisory Group (the ‘Advisory Group’) comprising:

- Professor Peter Klinken (WA Chief Scientist), Chair;
- Professor Fiona Stanley AC (Patron and Founding Director of the Telethon Kids Institute and Distinguished Research Professor of The University of Western Australia); and
- Mr Giles Nunis (Government Chief Information Officer).

Professor Bruce Armstrong (Emeritus Professor, School of Public Health of the University of Sydney), and Professor Louisa Jorm (Foundation Director, Centre for Big Data Research in Health of the University of New South Wales) provided advice to the Advisory Group.

The report’s findings and recommendations have been formed through a rigorous process of consultation. Firstly, through a call for submissions from Government, academia, research institutes and the not-for-profit sector. Secondly, via a series of detailed discussions with various stakeholders across these sectors. Thirdly, through engagement with the Director General of the Department of Health (DoH), Dr David Russell-Weisz and various staff members from the Data Linkage Branch (DLB) of the DoH.

WA has a long and highly successful history of linking data, dating back to the 1970s. Data were linked for research purposes in an attempt to identify trends and gain insights into human disease. Over the years, the WA Data Linkage System (WADLS) developed through a collaborative venture between the DoH, Telethon Kids Institute (TKI), Curtin University and The University of Western Australia (UWA).

The WADLS is administered by the DLB and delivers comprehensive, population-based data linkage. It is accessed predominantly by researchers aiming to identify links between disease and population data.
The WADLS has contributed to a range of policy, legislation and investment measures aiming to improve the health and wellbeing of Western Australians. This system has become internationally-renowned for its data linkage capability and has been the envy of the rest of the country.

The success of linking health data has attracted more than $136 million in research and related funding into the State from external sources. It has also supported over 400 studies, some of which relate to the improvement of mental health legislation, reducing criminal recidivism and changing vaccination schedules for children in the Kimberley region.

Linking health data has been so successful that a growing demand to link non-health data has emerged. This includes data relating to the justice and corrections system, disability services, community development, as well as training and education. Broadening data linkage beyond health is seen as an important opportunity if WA is to make the best data-driven policy decisions for the community, through a whole-of-Government approach.

However, the growing demand has sparked concern that the current method of delivering data linkage services by the DLB is not able to cope with a whole-of-Government approach. There is evidence to suggest that the current data linkage service is struggling with demand across the health sector, culminating in increased costs and longer wait times for users.

This review sought to explore these concerns so that WA’s data linkage capability is able to maintain its current high standards. In addition, the need to develop a whole-of-Government model required evaluation.

Importantly, the DLB should be commended on its ability to deliver a data linkage service with no breaches of privacy over several decades. The DLB has endeavoured to keep up with the growing demand for its services. Researchers, in particular, rely heavily on linked data and the work of the DLB, which should be well supported to deliver an efficient service. However, issues relating to factors such as governance, timeliness, transparency and costs have been raised as critical concerns and have been addressed in this report.

This report recognises the challenges faced by the DLB and has identified a number of recommendations that can be made in the short and medium-term, to ensure the growing demand can be better satisfied.
State Government plays a critical role for the future of data linkage. In the short-term, the review sees value in establishing a Steering Committee that provides strategic leadership on data linkage across Government. In addition, the Steering Committee should oversee the implementation of recommendations from this report.

Funding and resourcing by State Government and relevant agencies will be critical to ensure all recommendations are addressed, especially for the DoH in the short-term.

Increased engagement between the State Government and the not-for-profit sector is seen as an important short-term recommendation to enhance the State’s data linkage capability.

The protection of data used in the linking process is paramount to ensure high integrity and standards. Privacy legislation is an instrument that most jurisdictions across Australia have developed to protect an individual’s data. However, the review has found that even though WA does not have privacy legislation, this has not impeded the current services provided by the DLB with no breaches of privacy to date.

Nevertheless, there is evidence to suggest that other States and countries are hesitant to share data with WA due to the lack of privacy legislation. With this concern, coupled with the requirement for a whole-of-Government model where data from numerous Government agencies can be accessed and linked, there is a strong need for privacy legislation. The review has, therefore, recommended that the State Government draft privacy legislation and consider the formulation of data sharing legislation.

The Productivity Commission’s current inquiry into Data Availability and Use across Australia has identified a number of draft, nationally-consistent, recommendations for improving the availability and use of public and private sector data. If adopted, the recommendations may have some bearing on the recommendations and outcomes of this review – this will need to be assessed following finalisation of the inquiry in 2017.

The Office of the Government Chief Information Officer (OGCIO) will play a critical role in ensuring Government data is collected and managed in a consistent manner, so that data linkage services are delivered efficiently. It is essential that Government employees are supported through training and awareness initiatives relating to proper data management practices, and that the benefits of data linkage are communicated effectively. Building data analytics capability within Government will be necessary if the State is to embrace a whole-of-Government data linkage model.
In the longer-term, the review concluded that State Government establish a Statutory Body responsible for whole-of-Government data linkage. The Statutory Body would be overseen by a Governing Board which also includes members who are independent of Government and reflect wider community interests in data linkage.

In addition to the Statutory Body, the review also found that an ethics committee should be established to focus solely on data linkage. This ethics committee would have responsibility for the privacy and ethics approvals processes relating to data linkage, whether for research purposes or not, and must be compliant with the National Statement on Ethical Conduct in Human Research (NSECHR).

The findings and recommendations contained within this report will enable WA to improve on its current data linkage capabilities, and with time, deliver a whole-of-Government model that allows the linkage of health and non-health data. The information generated by data linkage will contribute immensely to research and Government policy decision-making, enabling further improvement to the health and wellbeing of Western Australians.
Recommendations

**Short Term**

*Department of Health/Data Linkage Branch*

1. The Data Linkage Branch be more open and transparent in all its dealings with stakeholders and clients, including about its policies, decision-making, timelines and charges. To support this, the Department of Health should establish a group (e.g. a working group, board or committee), similar to the Data Linkage Advisory Board, but with consideration to expanded membership.

2. The Data Linkage Branch, and any future linkage facility, continues to apply the current procedures for ensuring data security and privacy, and update them as required.

3. The Department of Health Human Research Ethics Committee, and any future Human Research Ethics Committee considering applications for data linkage, be flexible in its scope and operation. For example, it should consider larger programs of work, as well as discrete projects, and enable ongoing access to linked data sets.

4. The Data Linkage Branch, in conjunction with key stakeholders, develop and publish key milestones, and their corresponding anticipated timeframes, for delivering data linkage services.

5. The Data Linkage Branch develop an automated online linkage project application and tracking system. Up-to-date information on application status, milestones, actions required and anticipated completion dates should be included in this system and be fully available to applicants.

6. The Data Linkage Branch continue delivering client services to applicants and stakeholders, including providing feedback on draft applications for data linkage services.

7. The Data Linkage Branch provide additional guidance to clients, particularly those new to the service, through training and more fully annotated application documents. Similarly, Government and research organisations implement procedures to ensure appropriate quality of applications for linked data.
8. The Data Linkage Branch be more transparent regarding its charging for services, including: making its charging schedule and algorithms publicly available; increasing the level of detail provided in invoices; and providing justification to clients where the final charge is more than 10% above the final quote. The Data Linkage Branch should also review its charging schedule to ensure its percentage of cost recovery is consistent for all project types.

**Short Term**

**Whole-of-Government**

9. State Government, led by the Department of the Premier and Cabinet, establish a Steering Committee (Data Linkage Steering Committee) that provides strategic leadership and advice on data linkage and oversees implementation of the report recommendations.

10. The Department of Health, and all other State Government agencies, review and amend policies and procedures so that data custodians do not consider privacy and ethical issues (which includes research merit) when approving a data linkage application. This responsibility rests with the Human Research Ethics Committee.

11. Greater awareness of the benefits of sharing data is required across the State Government to improve the availability of data for data linkage.

12. That the Office of the Government Chief Information Officer develop a set of whole-of-Government standards or guidelines for the collection of data across Government, and facilitate training and support to public sector employees. All State Government agencies should be required to follow these guidelines and will need to make appropriate resources available.

13. The State Government improve its engagement with the not-for-profit sector to enable the sector to participate better in data linkage and analysis.

14. The State Government increase funding to the Data Linkage Branch. Greater resourcing is essential to support and improve data linkage service provision.
Medium Term

Department of Health/Data Linkage Branch

15. The Data Linkage Branch, through the Department of Health, and in conjunction with other State Government agencies, invest in increasing the number of data sets that use the Custodian Administered Research Extract Server.

16. The Data Linkage Branch, in conjunction with key stakeholders, pilot the use of data repositories for commonly linked and used data sets. The potential for greater use of data repositories should be evaluated as part of this project.

Medium Term

Whole-of-Government

17. The Department of the Premier and Cabinet and the Office of the Government Chief Information Officer seek approval from Cabinet to draft privacy legislation for Western Australia.

18. The Department of Premier and Cabinet and the Office of the Government Chief Information Officer investigate the feasibility of developing policy or legislative mechanisms to better enable the sharing and linking of State Government data.

19. State Government agencies are encouraged to explore the feasibility of accessing data collected by third parties e.g. service providers contracted by Government and other private sector organisations. Care would need to be taken if considering outsourcing data collection to the private sector, particularly with respect to ensuring data quality, value for money, and full and continued access to the data by Government.

20. State Government agencies invest in data analytics capabilities, and associated training courses for public sector staff be developed.

21. The Department of the Premier and Cabinet establish and oversee a unit which coordinates, between Government and the academic sector, cross-agency or whole-of-Government data linkage projects for policy development and implementation.
**Longer Term**

*Whole-of-Government*

22. The Department of Health, the Department of the Premier and Cabinet and the Office of the Government Chief Information Officer seek approval from Cabinet to draft legislation to create a Statutory Body that would take over the responsibility for providing future data linkage services for the whole-of-Government. This body should be overseen by a Governing Board, which also includes members who are independent of Government and reflect wider community interests in data linkage.

23. In conjunction with establishing the data linkage Statutory Body, a Human Research Ethics Committee with specialist expertise in data linkage should be created. The Human Research Ethics Committee should have responsibility for the privacy and ethics approval processes for data linkage, whether or not done for research purposes, and be compliant with the National Statement on Ethical Conduct in Human Research.
PART 1 – INTRODUCTION AND BACKGROUND
Chapter 1 - Introduction

A *Science Statement for Western Australia: Growing Western Australia* (the ‘Science Statement’) was launched in April 2015 by the Premier and Minister for Science, the Hon Colin Barnett MLA. The Science Statement highlights ‘Medicine and Health’ as a key priority for the State with WA being well placed to build on its internationally-recognised status in medical and health research. The Science Statement also highlights ‘Data’ as a cross-cutting area of opportunity that underpins all science priorities, citing the benefits of incorporating world’s best practice into the State’s data linkage system, leading to improved capacity for cutting-edge medical research to benefit the health and wellbeing of the community.

Further to this, the Science Statement outlines the following key priority for the Office of Science to achieve over a two year period:

> “Work with Government and other organisations to develop a new data linkage model that builds on the pioneering population level data linkage system in Western Australia’s health sector”.

This priority recognises and reinforces WA’s strengths and capabilities in data linkage and status as having a world-renowned data linkage system. The priority also seeks to build this capability and broaden the benefits to more than health outcomes through a whole-of-Government model that delivers outcomes to non-health areas such as education, child protection, corrections and social justice.

To achieve this priority, the Office of Science embarked on a project to investigate the current data linkage environment to understand:

- the needs and expectations of stakeholders and users;
- its history;
- case studies;
- benefits;
- governance arrangements; and
- the level of appetite across the sector of Government, academia and research institutes to broaden the current health data linkage model to a whole-of-Government approach.
The insights gathered through this process indicated the critical role of data linkage to deliver improved outcomes in health, with an emerging and exciting potential for data linkage to be applied to non-health data.

Over the years there have been numerous examples where data linkage has delivered economic benefit to WA, from understanding disease, making more informed decisions regarding health policy and predicting burden and associated costs to the health system for the best use of taxpayer dollars. The ways in which the lives of Western Australians have also been improved through data linkage will be explored throughout this report.

From these investigations it was clear that WA has a great history in data linkage with many initiatives delivered, and outcomes produced, for the benefit of the State. It is also evident that the current data linkage system located within the DoH should be commended for upholding privacy provisions that have translated to no breaches over its history. However, currently there appears to be challenges presented for the DoH in meeting a higher demand for data linkage in a timely and cost efficient manner. These challenges, and others, will be outlined in further detail in this report.

Following initial analysis of the data linkage environment within WA, it was deemed prudent to conduct an independent review into the State’s data linkage capability to ensure both a transparent and informed approach to the development of a whole-of-Government model as well as immediate actions to improve on the current system operating within the DoH. Without the review, and corresponding actions, there is the potential for the following to be compromised:

- retention of a highly esteemed, world-renowned data linkage capability;
- having a health system able to project demand for services and associated costs to maximise valuable taxpayer dollars;
- ability to understand why diseases occur in certain pockets of the population;
- ability to transfer the current data linkage capability to other policy settings; and
- being able to evaluate whole-of-Government policy.
On 17 March 2016 the review was authorised by Mr Peter Conran, the then Director General of DPC. Endorsement of the review was also provided by Dr David Russell-Weisz, the Director General of the DoH. The Hon Colin Barnett MLA, Premier and Minister for Science, announced the review on 3 May 2016.

The Advisory Group was established to undertake the review and deliver on the Terms of Reference (ToR) (Appendix 1). The Advisory Group membership included Professor Peter Klinken (WA Chief Scientist, as Chair) Professor Fiona Stanley AC (Patron and Founding Director, Telethon Kids Institute and Distinguished Research Professor, The University of Western Australia) and Mr Giles Nunis (WA Government Chief Information Officer). Assisting members were Professor Bruce Armstrong (Emeritus Professor at the University of Sydney’s School of Public Health) and Professor Louisa Jorm (Foundation Director of University of New South Wales’ Centre for Big Data Research in Health). Support for the Advisory Group was provided by the Office of Science, DPC. Short profiles of the members are included at Appendix 2.

Subsequent to the Premier’s announcement of the review on 3 May 2016, a call for submissions commenced on 16 May 2016 with a closing date of 30 June 2016. On 7 June 2016 an information session about the review, and related processes, was held by the Advisory Group at Dumas House, 2 Havelock Street, West Perth. More than 70 people from across Government, research, industry and the not-for-profit sector attended. A list of attendees is provided at Appendix 3. Mr Peter Conran opened the session and Professor Peter Klinken spoke about the review on behalf of the Advisory Group. All in attendance were provided with an opportunity to ask questions regarding the review and its process.

During the information session, guest presenters Professor Desiree Silva (Head of Paediatrics, Joondalup Health Campus) and Mr Damian Shepherd (Director of the WA Land Information System Office, Landgate) demonstrated the broad applications and benefits of data linkage.

Following the call for submissions, a total of 41 submissions were received, comprising sixteen from Government agencies (fifteen from the WA Government and one from the NSW Government), twenty from the research and academic sectors, three individual submissions and two from the not-for-profit sector.

---

Approximately half of the submissions were from the health or health-related sector. A list of submissions received is provided at Appendix 4.

Further to the receipt of submissions, the Advisory Group held a number of meetings with a range of stakeholders to discuss their submissions in greater detail. A list of the meetings held is provided at Appendix 5. The submissions received and meetings held provided valuable information to the review conducted by the Advisory Group.
Chapter 2 – Background

2.1. Definition of data linkage

Data linkage is a technique that connects pieces of information that are thought to relate to the same person, family, place or event. Connecting sets of data can provide a more complete picture of how certain factors interrelate and impact across a variety of domains. Data linkage can provide many insights, some of which can include: a) determining whether intervention programs are working or failing, b) identifying areas that need greater resources, c) tracking trends to promote novel intervention strategies, and d) gathering information regarding risk or preventive factors. Linking data across Government agencies and services can provide a powerful resource for understanding how complex issues can be solved across agencies, and aids the development of new multidisciplinary approaches to improve services and outcomes for the WA community.

Over recent years there has been a rapid increase in the demand for linked data for research and Government planning. Much of this is in response to a realisation that many “wicked problems” facing modern democracies require a whole-of-Government response (for example: social health determinants, climate change, inequalities, youth incarceration, mental health, youth disengagement and educational issues). A whole-of-Government data capability can enable a more informed approach to solving these problems and an ability to better evaluate whether current or proposed services and policies are working, or are likely to work.

Each person in the world creates a Book of Life. This Book starts with birth and ends with death. Its pages are made up of the records of the principal events in life. Record linkage is the name given to the process of assembling the pages of this Book into a volume’

*Dunn HL. Record linkage. Am J Public Health (1946; 36: 1412-1416)*
2.2 Overview of the data linkage process in WA

Linking individual data sets creates the potential to better inform wide-scale decisions with far reaching impacts that can benefit communities and the economy. Researchers and other users rely heavily on linked data to be able to understand the interrelationships between people and the world so that the health and wellbeing of individuals can be improved.

The DLB of the DoH, together with a number of partners, has been a pioneer of data linkage in WA. The process involves a number of people and organisations navigating the data linkage framework in order to deliver linked data. Figure 1 below provides a broad overview of the data linkage framework in WA. More detail is provided at Appendix 6.

Figure 1: WA Data Linkage Framework – Overview

Data linkage cannot occur unless appropriate mechanisms for data collection are in place. Data is collected by a number of organisations including State and Commonwealth Government agencies, research institutes, not-for-profit organisations and industry. The DLB within the DoH governs and administers data linkage services using WA Government data and other data sets.
A person or team (the ‘applicant’) seeking access to linked data must first prepare a justification and submit an application for data to be linked to the DLB. All applications are reviewed and approved by the relevant data custodians (the person within an organisation or agency that is responsible for the appropriate collection and management of the data) and also by the DoH Human Research Ethics Committee (HREC). Applicants are required to adhere to a number of policies relating to use of information, access and disclosure, governance and ethical review procedures prior to gaining access to the data. Once the data custodians, the DLB and the HREC are satisfied that the research is appropriate and data linkage is possible, the data is linked and released to the applicant.

2.3 History of data linkage in WA

Understanding the history of data linkage in WA, and the strength and commitment taken to deliver a world-renowned data linkage capability, will be critical to setting the path forward.

In WA, data linkage started in the late 1970s where all deaths, hospitalisations and mental health service data sets dating back to 1966 were captured, linked by the DoH and analysed by public health academics in the Department of Medicine at UWA.

By the 1980s, new databases and improved data analysis methods were spreading. This period saw the establishment of the WA Maternal and Child Health Research Database and of the WA Land Information System (WALIS).

The WA Maternal and Child Health Research Database was created by researchers in UWA’s National Health and Medical Research Council (NHMRC)-supported Research Unit in Epidemiology and Preventative Medicine (which later formed the nucleus of the Telethon Institute for Child Health Research/TKI). This database linked data from births, deaths, midwives notifications of births, and hospitalisations to perform epidemiological, paediatric and perinatal studies. Registers of birth defects and cerebral palsies were established and linked into this database. This enabled analyses of pathways from conceptions into outcomes and to monitor the impact of these disabilities.

In 1981 the WA Government formally adopted data linkage for its land information systems. WALIS emerged as an initiative to improve efficiency within the location and spatial data industry by reducing duplication and improving access to the rapidly growing demand for land tenure and cadastral information.
Since this time, the Shared Land Information Platform (SLIP) was developed from the WALIS concept, allowing users to access Government’s significant land and geographic information resources. Leveraging off SLIP, SPUR has emerged as an innovation hub to promote collaboration, stimulation and acceleration of new ideas through accessing data, resources and key contacts.

In 1995 the WA Lotteries Commission awarded an infrastructure grant to UWA to deliver a comprehensive, population-based data linkage system in WA. Through this, WADLS was established and represented a strong collaboration between the DoH, UWA, Curtin University and TKI. The DLB was established within the DoH and delivered data linkage services through WADLS.

The DLB has connected a range of available health and related information for the WA population, for use in ethically-approved research, planning and evaluation aimed at improving the health of Western Australians. Developed under the stewardship of Government, academic and consumer organisations, WADLS is one of the most comprehensive linkage systems available in the world.

A peer review undertaken in 1999 placed WADLS as one of only five long-term comprehensive record linkage projects in the world. At that time, more than eight million health records dating back to 1980 had been linked from six core data sets, mostly held by the DoH, including:

- birth records;
- midwives’ notifications;
- cancer registrations;
- inpatient hospital morbidity;
- inpatient and public outpatient mental health services; and
- death registration.

With some data sets spanning back more than 45 years, and linked information for a total historic population of 3.7 million people, the depth and breadth of WADLS is almost unrivalled world-wide.
Over the years, WADLS continued to expand, initially from a linkage key to local health data sets, to now encompassing links to national and local health and welfare data sets, genealogical links and spatial references for mapping applications. See Figure 2 which shows the status of WADLS data linkage as at November 2016.

Figure 2: WADLS status of data linkage as at November 2016

Projects using the WADLS have contributed to a range of policy, legislation and investment measures improving the health and wellbeing of Western Australians. During the first ten years of full operation, WADLS supported over 400 studies with more than 250 journal publications and 35 graduate research degrees. WADLS has attracted more than $136 million in research and related funding into the State from external sources.

---


3 'Chair in Public Health: 20 Years of Discovery and Advance 1994-95 to 2013-14', 2014. School of Population Health, The University of Western Australia, pp 31
More specific examples of the benefits and outcomes of WADLS are included toward the end of this Chapter. Figure 3 below shows the data sets linked through WADLS and the calendar periods of data that have been linked.

Figure 3: Data sets linked through WADLS

WADLS was further expanded from 2005 through the establishment of the Developmental Pathways Project (DPP) at TKI in collaboration with the DoH and three other State Government agencies. The DPP now has fourteen partnering State Government agencies from across the health and social services sector.
The purpose of the DPP was to expand WADLS to include other social services data such as education, justice and disability, to build a more comprehensive understanding of the interconnectedness of factors affecting the health and social outcomes of WA children and youth. Some of the examples of the benefits and outcomes of data linkage featured in this Chapter are from the work of the DPP.

The three-tiered governance structure of the DPP includes representatives from all of the partner organisations, and at a range of levels. It has been funded through two consecutive Australian Research Council (ARC) linkage grants and contributions from the partners.

Also in 2005, the WA Government Centre of Excellence Program (CoE) provided funding to establish a multi-party CoE for Data Linkage Australia to expand the capacity of data linkage in the State, including upgraded hardware, system programmers and the development of national and international collaborations. The four core partners were the DoH, Curtin University, UWA and TKI.

At conclusion of the State Government’s CoE funding period in 2011, the partners agreed to approach Lotterywest for support to continue the data linkage infrastructure development program. The aim of the Lotterywest WA Data Linkage Infrastructure Project (WADLIP) was to strategically upgrade and strengthen WA’s data linkage infrastructure to support health and medical research, and provide opportunities for collaboration with other national and international data linkage centres. In February 2013, Lotterywest approved $5.1 million for WADLIP, to conclude in mid-2017. A Governance Board comprising the partners was established to oversee the use of the funding.

The DLB has indicated that while grant funding is of assistance to its activities, it is often time limited and allocated for a specific purpose.4

In 2009, the Population Health Research Network (PHRN) was established with $20 million over four years from the Commonwealth Government’s National Collaborative Research Infrastructure Strategy (NCRIS) program.5 The PHRN aims to “build a nationwide data linkage infrastructure” and is made up of a network of data linkage units throughout Australia.

---

4 2016 Data Linkage Review – Submission Process
WA plays a prominent role in this national network, with the national program office run by UWA and the DLB acting as a node of PHRN. Also, the Centre for Data Linkage (CDL) was established at Curtin University and is one of PHRN’s main data linkage facilities for national projects. The CDL is responsible for enabling linkage between jurisdictional data sets, and between these data sets and research data sets for the purpose of health and health-related research projects.

The expertise and sophistication of WA’s data linkage system paved the way for the establishment of many of the other State’s data linkage systems and nodes – more information is provided later in this Chapter.

The review of WA data linkage capabilities was announced in 2016 to investigate the current data linkage environment in WA and map a path to improve and expand existing systems. Looking forward, it is expected that data linkage will play an increasingly important role in improving outcomes for Western Australians for generations to come.

2.4 National data linkage landscape and initiatives

**PHRN and other State and Territory Data Linkage Units**

- The NSW Centre for Health Record Linkage (CHeReL) was established in 2006 and is managed by the NSW Ministry of Health. The manager of CHeReL reports to the Deputy Secretary, Population and Public Health and Chief Health Officer. Organisations may become members of CHeReL and participate in the Data Linkage Advisory Committee, which provides advice to the NSW Ministry of Health on CHeReL’s strategic plan, the development of record linkage infrastructure and services, and the use of linked data to enhance policy-relevant research.

- HealthLinQ was established in Queensland (QLD) in 2009 and is managed by the QLD Department of Health (DoH). The centre facilitates access to linked administrative data and provides consultation services to researchers interested in using linked administrative health data. HealthLinQ was previously managed by the University of Queensland but was transferred to the Health Statistics Unit in the QLD DoH in June 2014.
SA-NT DataLink was established in South Australia (SA) and the Northern Territory (NT) in 2009 and provides funding and in-kind contributions for the operation of the data linkage in those States. SA-NT DataLink operates on a consortium model (unincorporated joint venture) governed by a Steering Committee as described in the SA NT DataLink Consortium Agreement, with members comprising the SA Government, the NT Government, various universities, research institutes and others. Members contribute funding to DataLink and help govern the process. SA-NT Datalink’s head office is located in the South Australian Health and Medical Research Institute.

The Centre for Victorian Data Linkage (CVDL) was established in 2009 (previously known as the Victorian Data Linkage Unit). CVDL is managed by the Victorian (VIC) Department of Health and Human Services and is supported by the VIC Department of State Development, Business and Innovation. CVDL is focused on data collection supporting health and wellbeing policy and research, overlooks the efficient and accurate linkage of data, and facilitates the use and access of linked data in VIC.

The Tasmanian Data Linkage Unit (TDLU) was established in Tasmania (TAS) in 2010 and is operated by the Menzies Institute for Medical Research under a contractual arrangement between the TAS Department of Health and Human Services and the University of Tasmania. The TDLU creates links between data relating to individuals in TAS across multiple administrative and clinical data sets from various sectors and supports nationally and internationally significant population-based research with a focus on improving health outcomes.

Other national organisations and initiatives
Other national institutions engaged in data linkage include:

- Australian Institute of Health and Welfare (AIHW) – provides data linkage services ranging from routine linkage between their data sets (i.e. National Death Index and Australian Cancer Database) to more complex data integration with sets from other Government institutions;
- Australian Bureau of Statistics (ABS) – heavily involved in data linkage (referred to as “data integration” within the ABS) initiatives for a range of statistical and research purposes;
- Australian Institute of Family Studies (AIFS) – AIFS, like AIHW and ABS, is an accredited Data Integrating Authority, meaning the Institute is authorised to undertake data integration projects involving Commonwealth data for statistical and research purposes;
• The Sax Institute – a not-for-profit organisation that promotes research backed by data in a number of health policy initiatives, including SURE and active participation in PHRN; and
• CSIRO’s Data61 – an entity created between CSIRO and National ICT Australia aimed at connecting disparate Government data sets to deliver economic, innovation and industry benefits to Australia.

Other recent national data initiatives demonstrate the growing interest in utilising the power of “big” and linked data. These include:

• The Productivity Commission’s Public Inquiry into Data Availability and Use, currently in progress, has identified a number of draft, nationally-consistent, recommendations for improving the availability and use of public and private sector data. The final inquiry report is expected to be handed to the Australian Government in March 2017;
• The Department of the Prime Minister and Cabinet’s (DPMC) Public Sector Data Management Project and establishment of data.gov.au as an open data platform;
• The ABS’ Transforming Statistics Program, which includes working with the Australian Tax Office (ATO) and others, and the Multi-Agency Data Integration Project (MADIP), a collaborative cross-portfolio partnership to improve accessibility to public data;
• The Council of Australian Governments (COAG) is pursuing initiatives that will enhance transparency and provide Australian citizens with a greater level of real time data on government expenditure, outcomes and performance; and
• The National Disability Insurance Scheme (NDIS) is envisioned to make full use of statistical data linkage across all the disability data sets made as the program progresses.
2.5 Benefits of data linkage

The following section presents case studies that demonstrate the benefits of data linkage to Western Australians and provides a rationale as to why the current system should be protected and expanded.

**Child maltreatment and early intervention**

Data linkage relating to child maltreatment has been instrumental in guiding the Department for Child Protection and Family Support’s (DCPFS) ‘Building Safe and Strong Families – Early Intervention and Family Support Strategy’.

A number of research studies using cross-departmental data linkage between DCPFS, DoH, the Mental Health Commission and the Disability Services Commission identified the primary drivers for child protection interventions in WA, including family and domestic violence, parental substance abuse, and parental mental health issues.

Early intervention can bring significant socio-economic benefits to WA. Approximately 1,398,735 total nights, at a cost of $176 per night, of out-of-home care were registered in child protection services in 2014-15. It is estimated that each 1% reduction in the need for out-of-home care would save $2.5 million per annum.

---

6 Including:
**SLIP and the South-West Native Title Settlement**

Thousands of datasets have been packaged together by Landgate and its partners in SLIP to make it easy to bring together location information to view and share through maps or use in geographical information systems and other business applications.

SLIP is part of the WA Open Data service at data.wa.gov.au and has improved access to, and use of, location-based data sharing and linkage. Benefits realised to date include significant savings to the State via synergies, more efficient data sharing between the public and private sectors and improved citizen engagement via interactive mapping services. In 2015-16 alone, Landgate delivered 50 new SLIP services to customers across the public, private and research sectors.

For example, new SLIP services were used to support the South-West Native Title Settlement, the most comprehensive native title agreement in Australian history. This historic agreement involved around 30,000 Noongar people and covered about 200,000 square kilometres. SLIP enabled the Department of Lands, the South-West Aboriginal Land and Sea Council and the Department of Mines and Petroleum a shared view of the same up-to-date information. This significantly reduced the time and effort involved in exchanging and analysing interests in Crown land being negotiated as part of the agreement.⁸

**Emergency management**

The WALIS framework in WA has helped to ensure that WA Government agencies involved in emergency management have access to the available foundation data from relevant agencies and data custodians. SLIP provides a means to access that data. The establishment of web mapping services across Australia and internationally is assisting better sharing and real time access to information. For example, the Department of Fire and Emergency Services, and others, use these web services to access real time weather information from the Bureau of Meteorology and real time satellite monitoring services from external organisations, and to share operational information between agencies (some of this through SLIP).

---

**Improving mental health legislation**

A study using WADLS enabled the first comprehensive picture of the overall health of people with mental illness and highlighted a significantly increased rate of serious physical illness in mentally ill people, often resulting in death\(^9\). The study’s findings were published in 2001 and led to a number of changes, including amendments to the *Mental Health Act 1996* (WA) to include compulsory physical examinations of psychiatric inpatients by medical practitioners and discharge planning that covered physical as well as mental health. A $20 million State Government investment into health services for people with mental illness in regional WA was also committed as a result.

**Improving outcomes for children with attention deficit hyperactivity disorder (ADHD)**

A study using linked data found that children treated for ADHD are more likely to come into contact with the juvenile justice system and commit offences at a younger age\(^10\). Early diagnosis and coordinated management of ADHD may reduce the over-representation of children and young people with the disorder in the juvenile justice system and reduce the likelihood of serious criminal activity and reoffending in the future. The use of linked education and police data may assist with early detection of ADHD and facilitate intervention strategies aimed at reducing first contact with the justice system.

**Informing strategic health policy**

Data linkage has been of great benefit in projecting the estimated costs of health system services in WA. One example is the cost impact on the health system of fall-related injuries. In one study it was estimated that the cost of falls in 2004 to the health system was $86.4 million with a projected cost to be $181 million in 2021 should no intervention occur\(^11\). This projection was made possible through data linkage and enabled the health system to make long-term strategic policy decisions on the prevention of falls.

---

\(^9\) Coghlan, R, Lawrence, D, Holman d, Jablensky A, 2001, ‘Duty to Care: Physical illness in people with mental illness’. Perth: The University of Western Australia


Also a 2006 study\textsuperscript{12} found many of the high-cost users of the hospital system in WA had chronic conditions (such as cancer, end-stage renal disease and congestive heart failure), that have identifiable and modifiable risk factors (such as type 2 diabetes, obesity, tobacco use and alcohol consumption). This highlighted the potential to reduce medical costs in the State through public health campaigns and strategies aiming to prevent the onset and limit the progression of some types of chronic disease.

Being able to use linked data to project cost impacts in the health system paves the way for further research into intervention strategies that contribute to easing the burden on the State’s health system, thus maximising use of taxpayer dollars.

As another example, it is estimated that each 1% reduction in emergency room presentations would save $6.3 million per annum\textsuperscript{13}.

\textit{Reducing recidivism}

Linked data research has been used in WA to better understand and possibly reduce recidivism in the State. UWA’s Crime Research Centre was operational between 1993 and 2014 as a collaboration between UWA, WA Police, the Department of the Attorney General and the Department of Corrective Services. The Centre’s work led to the development of the Integrated Numerical Offender Identification System, which facilitated the creation of the de-identified, linked Crime Research Centre Offender Database. The system enabled a variety of useful research including studies about recidivism, analysis of criminal justice processes and evaluations of criminal justice interventions and programs.\textsuperscript{14}

Linking data can aid the State with targeted intervention efforts aimed at reducing recidivism rates – one study reported that non-prison interventions for youth offenders resulted in a recidivism rate of 31.5% versus 41.3% for the traditional prison system\textsuperscript{15}.

Furthermore, with each prisoner costing WA taxpayers an estimated $360 dollars per prisoner-day, reducing contact with the criminal justice system would translate into significant savings for the State. It is estimated that each 1% reduction in prisoner-days in adult corrective services saves $7.1 million per annum, each 1% reduction in the number of criminal matters in courts saves $1.3 million per annum and each 1% reduction in youth detention saves $0.5 million per annum.\(^\text{16}\)

**Vaccinating children from the Kimberley region**

Through using linked hospital data with an individual’s region of residence, researchers determined that children living in regions of WA with a tropical climate have different patterns of hospitalisations for bronchiolitis (a viral infection of the lungs) than children in regions within temperate and sub-tropical climate zones.

The study found the disease to be highly seasonal, with clearly defined epidemic peaks in July or August in southern WA and no identifiable seasonal pattern for the Kimberley region.\(^\text{17}\) This research provided evidence supporting different immunisation campaign schedules throughout the State. Preventative programs are a valuable asset for WA’s health system. For example, it is estimated that each 1% reduction in emergency room presentations would save $6.3 million per annum.\(^\text{18}\)

### 2.6 International examples

It is important to understand the global environment for WA to remain at the forefront of data linkage. The information below provides a brief outline of several other noted data linkage systems around the world. While they all aim to use linked data to produce better health and social outcomes for the public, they all have differing systems and processes due to divergent end users. These systems are most effective when tailored to their respective audience. Although different, these systems work to produce data linkage infrastructure suitable for each jurisdiction’s specific legislation and aims; there is no “one model fits all” solution.


\(^\text{18}\) See note 16 above
**New Zealand**

The New Zealand Integrated Data Infrastructure (IDI) aims to use whole-of-Government data analytics to produce better policies and outcomes for the public. It was established in 2012, with Statistics New Zealand (Statistics NZ), a Government organisation, being able to provide data integration services. It is now a central data management framework managing social and health data from various partner agencies including the Ministry of Health, Ministry of Justice and the Ministry of Social Development. The IDI collects this data from various Government and non-Government organisations, with the majority of data updated every quarter and linked in-house. Statistics NZ is led by its Chief Executive, the NZ Government Statistician, and reports to the Minister of Statistics.

The IDI has stringent privacy and security processes in place to ensure that data is accessed and used appropriately. In order for applicants to access data, they must meet a number of conditions:

- Researchers must pass referee checks and sign a declaration of secrecy under the *Statistics Act 1975* to ensure that rules and protocols are followed;
- Researchers must prove that their research is in the public interest and the Government Statistician must sign off on all research proposals;
- A secure Data Lab must be used to access the data where computers are not connected to a network and only Statistics NZ staff can release the de-identified data specific to the research project; and
- Any results from the research must be checked by Statistics NZ staff before it is released, to ensure that any risk of re-identification of individuals from the results is negligible.\(^\text{19}\)

**Canada**

Statistics Canada is the main organisation for linked social science and health survey data in Canada. It is Canada’s central statistical office and acts as the main access point for external data users as well as providing its own in-house data analysis and custom record linkage. In order to access this data, researchers must request access through a Research Data Centre or a remote access facility, which must then be approved by their funding agency and Statistics Canada.

Through this approval, the researcher is made a “deemed employee” of Statistics Canada and is able to access linked data and create new linkages once approved by the senior management of Statistics Canada, along with other ethics approvals.

The data from Statistics Canada is often linked to de-identified social and health data held by the University of Manitoba, which hosts the province of Manitoba’s Centre for Health Policy (MCHP). The mission of MCHP is to conduct research on the health of the people of Manitoba and provide accurate and timely information to healthcare decision-makers. Similarly to the data linkage program in Wales described below, the data custodian and trusted third party are within Government.

That is, Manitoba Health, a department within the Government of Manitoba, performs the linkage and the de-identified database (The Population Health Research Data Repository) of health, education and human services data is maintained by the University of Manitoba. The linked data can be accessed by accredited researchers on site at the University or via a number of Remote Access Sites.

Population Data BC is a health data resource network for the Canadian province of British Colombia. The University of British Colombia is responsible for the linkage, storage and maintenance of health data within the network. A Secure Remote Environment (SRE) ensures that identifiable data and content data are stored separately in order to provide a secure data analysis environment.

**United Kingdom**

The Farr Institute was established in 2013 after a consortium of ten funders, along with the Medical Research Council, which contributed a total of £37.5 million to create a world-class infrastructure for research using “big data”. It was formed as a research collaboration between twenty-one academic institutions and Government across England, Scotland and Wales, as well as other partners across the academic, Government and industry sectors. It aims to deliver cutting-edge research using data to improve the health of the public. The Farr Institute does not control data, but instead has set up four centres that produce some of the linked data used to carry out research, with other data obtained from various other partner agencies and organisations.
**Wales**

The Secure Anonymised Information Linkage Databank (SAIL) in Wales is managed and maintained by the Health Information Research Unit (HIRU). The HIRU manages and maintains the non-identifiable health and social data stored at the University of Swansea. Separate to this unit, the National Health Service (NHS) Wales Informatics Service acts as the data custodian and trusted third-party. By separating the linkage function performed by NHS Wales and the storage and maintenance of the non-identifiable data by the HIRU within the SAIL, it ensures the privacy and security of the data.

**Scotland**

The Scottish Informatics Programme (SHIP) is a collaboration across the academic and Government sectors. Now transferred to the Farr Institute of Health Informatics Research (see above) SHIP facilitates linkage between third party data sets and Government collected health data, but does not store or maintain these data collections. Projects must first be approved before data sets are released to an indexing service and analysis can only be undertaken by researchers at a designated “safe haven” or at their own institution.
PART 2 – KEY ISSUES & RECOMMENDATIONS
Chapter 3 – Governance

RECOMMENDATIONS

• The Data Linkage Branch be more open and transparent in all its dealings with stakeholders and clients, including about its policies, decision-making, timelines and charges. To support this, the Department of Health should establish a group (e.g. a working group, board or committee), similar to the Data Linkage Advisory Board, but with consideration to expanded membership. (Recommendation 1)

• State Government, led by the Department of the Premier and Cabinet, establish a Steering Committee (Data Linkage Steering Committee) that provides strategic leadership and advice on data linkage and oversees implementation of the report recommendations. (Recommendation 9)

• The Department of Health, the Department of the Premier and Cabinet and the Office of the Government Chief Information Officer seek approval from Cabinet to draft legislation to create a Statutory Body that would take over the responsibility for providing future data linkage services for the whole-of-Government. This body should be overseen by a Governing Board, which also includes members who are independent of Government and reflect wider community interests in data linkage. (Recommendation 22)
3.1 Current governance arrangements

Currently, the governance and legal responsibility of WADLS and staff of the DLB rests with the DoH. This has remained the case since the DoH became the lead agency for the operational aspects of data linkage in 2007 when it subsequently established the DLB and the Data Linkage Advisory Board (DLAB). The DLAB was responsible for:

- strategic growth of the WA data linkage infrastructure;
- privacy considerations of linked health data;
- budget funding and review;
- IT infrastructure; and
- access to linked data and training needs.

The DLAB has not met since 2012 which raises concerns as to the current mechanisms that oversee the DLB. It is noted that all members of the DLAB are currently members of the WADLIP Board. Nevertheless, stakeholders have cited a number of concerns relating to the current governance arrangements, including: a) a cumbersome approach to governance, b) the lack of transparency of the governance structures, c) the lack of a high level governing structure which could improve the level of coordination, and d) the need for membership that includes whole-of-Government approaches and representation from consumers and the research community.  

FINDING 1
The current governance structure of the DLB, viz the DLAB, has not met since 2012. This indicates absence of a strong, formal governance arrangement that provides adequate openness, transparency, accountability and consultation surrounding the functions of the DLB.

3.2 Inclusiveness and consultation

There appear to be limited avenues by which stakeholders can contribute to, or influence, the decisions made by the DLB. Within Australia, linkage units have taken a variety of approaches to addressing the issue of user input. Of the six State data linkage units – the DLB (WA), DataLink (SA-NT), CHeReL (NSW), TDLU (TAS), HealthLinQ (QLD) and CVDL (VIC) – four sit within their respective Government health departments.

---

20 2016 Data Linkage Review – Submission Process
Of these four, three (including WA) have no formal mechanism for user or stakeholder input, while one (CHeReL in NSW) has an advisory committee that can make non-binding recommendations to the unit. TDLU is based within a university and SA-NT DataLink is a consortium model, led out of a research institute, that includes stakeholders from academia, Government and research institutes.

The level of user and other stakeholder involvement could be increased by reconvening a body with the functions of the DLAB and including a mixed membership that is representative of the DLB’s key stakeholders, e.g. Government agencies, research institutes, universities, other data users and community groups.

**FINDING 2**
A lack of user input to the data linkage process is raising concerns regarding the absence of a stronger, broader governance mechanism. Maintaining open and collegiate relationships with stakeholders will assist the DLB in maintaining its effectiveness and reputation.

### 3.3 Transparency

A number of submissions to this review expressed concerns over the lack of transparency of operations within the DLB. Most concerns were raised with regard to the lack of communication relating to timeframes of data linkage requests, charging policies and significant changes thereof, as well as various decision-making processes.

Transparency and communication are critical aspects of an effective and efficient data linkage service. This includes communication in relation to important changes made to processes that support the linkage service and key decisions that impact stakeholders. Transparency and open communication ensures a productive working relationship between users and data providers. This is usually achieved through a tighter governance arrangement where the linkage provider must ensure the best interests of its stakeholders are considered, through open channels of communication.

Given the large number of concerns relating to transparency, these issues have been addressed in Chapter 7 – *DLB Processes* and Chapter 8 – *Funding and Linkage Charges*.

**FINDING 3**
Clear communication and transparency is essential for the provision of a data linkage service to maintain positive relationships with its stakeholders.
**RECOMMENDATION**
The DLB be more open and transparent in all its dealings with stakeholders and clients, including about its policies, decision-making, timelines and charges. To support this, the Department of Health should establish a group (e.g. a working group, board or committee), similar to the Data Linkage Advisory Board, but with consideration to expanded membership. (Recommendation 1)

### 3.4 Future governance – a centralised approach

Government is becoming more cognisant of the value of data linkage for purposes that extend well beyond the current usage in medical research. The scope of work that is now taking place, through the DLB, in the social service sector is expanding with infrastructure linkages now including data sets such as school records, courts data and incarceration data.

Half of the DLB’s business currently supports the DoH’s operational needs, particularly with respect to performance reporting, and this is rapidly growing with the DoH looking to continue utilising data linkage functions for operational purposes. Moreover, the Department of the Attorney General, the WA Police and a number of other agencies are working on the creation of their own data sharing and linkage functions to serve their operational needs.

The DoH has significant experience in the management and conduct of data linkage for medical research; however, this does not extend to other types of research or use of linked data. A number of submissions to this review expressed concern regarding the current DoH oversight of the DLB, noting that the current model is health-centric or health-focussed and lacks strategic vision.

**FINDING 4**
As the scope and usage of data linkage continues to grow beyond the health context, it becomes less appropriate for the data linkage function to be delivered through the DoH.

Concern has also been raised by stakeholders regarding the cumbersome layers of governance. In addition, the lack of a high level governance structure that provides the strategic leadership and oversight required, especially for developing a future whole-of-Government model, was raised as a major issue.
FINDING 5
There is currently a lack of high-level governance for data linkage to deliver leadership, vision, guidance and accountability.

Therefore, a higher level governance structure will be required to provide the necessary leadership and oversight for data linkage in WA. The Advisory Group recognises that this oversight role will require additional commitment and resourcing from the State Government and relevant agencies.

RECOMMENDATION
State Government, led by the Department of the Premier and Cabinet, establish a Steering Committee (Data Linkage Steering Committee) that provides strategic leadership and advice on data linkage and oversees implementation of the report recommendations. (Recommendation 9)

Broad support for centralising the State’s data linkage capacity was presented to this review by researchers and Government agencies, with fourteen submissions advocating for it in some form. A number of potential locations for a whole-of-Government data unit were proposed:

- a Statutory Body;
- an Office or Unit within a centralised Government agency; or
- a university.

Statutory Body
A Statutory Body could be created to provide data linkage capacity. This body would be governed by a Board and receive direct funding from the State. Its sole purpose would be to perform data linkage for both Government and external researchers.

Advantages:
1. Independent and overseen by a Governing Board would allow representation from different disciplines, government agencies, and sectors. Funding would come directly from the Department of Treasury (Treasury).
2. Focus exclusively on data linkage. In contrast, a linkage unit within either a Government agency or a university may have other priorities and would have to compete for resources internally.
3. Funded in accordance with Government priorities and devote all its resources to its statutory functions.
4. Governed by a Board to reduce community concerns about the Government’s access to, and use of, citizens’ data.

Disadvantages:

1. Creation and upkeep of the Statutory Body would require additional funding. Currently, funding for data linkage is covered by the DoH; additional resources would be required to provide more data linkage capacity and establish and maintain a Governing Board, along with administrative costs.

2. Could not be created quickly. The preparation and passage of legislation necessary to establish the Statutory Body, together with the appointment of a suitable Governing Board, the potential need for Memoranda of Understanding with data-providing agencies and the implementation of the necessary policies and procedures for the unit to operate, could take several years.

**Office or Unit within a central agency**

A number of submissions to this review suggested that data linkage capacity be created within a central agency, with Treasury, the OGCIO and DPC suggested as potential locations.

Advantages:

1. Central agencies have existing governance structures in place into which a data linkage unit could be placed.

Disadvantages:

1. Central agencies are policy-focused and have little experience running fee-for-service operations. As a result the unit may not receive the necessary support.

2. If placed within Treasury there could be concern that the data linkage capacity might be utilised by the Department to focus primarily on cost-cutting measures.

3. If placed within the OGCIO this additional activity may overwhelm the OGCIO and distract the agency from its core business.

4. The DPC is a policy-based Department. Taking on a data linkage function would be outside the remit to support the Premier and Cabinet.
Within a university
It was noted that Universities could be a potential location for a data linkage capacity, based on the success with this model in some foreign jurisdictions such as Manitoba Centre for Health Policy, University of Manitoba, Canada and SAIL in Wales.

Advantages:
1. Universities have significant expertise in the use and manipulation of large data sets.
2. Such a capacity housed within a university could be expected to benefit from a steady stream of graduate students and researchers contributing to its operation.
3. Curtin University currently houses the CDL, which is one of two PHRN national data linkage units. In addition, both Curtin University and UWA were founding collaborators of the current WADLS.

Disadvantages:
1. Housing such a capacity outside of Government may raise concerns in the community, and within Government, about the security of the data.
2. The location of the data linkage unit outside of Government could limit the control Government would have.

Conclusion
The most appropriate location for a whole-of-Government data linkage capacity is within a Statutory Body. While this model may take some time to implement, it is the only option that will ensure broad governance, while maintaining sufficient State Government oversight of its operations.

RECOMMENDATION
The Department of Health, the Department of the Premier and Cabinet and the Office of the Government Chief Information Officer seek approval from Cabinet to draft legislation to create a Statutory Body that would take over the responsibility for providing future data linkage services for the whole-of-Government. This body should be overseen by a Governing Board, which also includes members who are independent of Government and reflect wider community interests in data linkage. (Recommendation 22)
Chapter 4 – Privacy

RECOMMENDATIONS

- The Data Linkage Branch, and any future linkage facility, continues to apply the current procedures for ensuring data security and privacy, and update them as required. (Recommendation 2)

- The Department of the Premier and Cabinet and the Office of the Government Chief Information Officer seek approval from Cabinet to draft privacy legislation for Western Australia. (Recommendation 17)
4.1 Privacy of personal information and data linkage

Privacy of personal information relates to how information/data is managed so that an individual cannot be identified.

Data linkage relies heavily on the use of an individual’s personal information. As a result, community support for data linkage requires confidence that the privacy of an individual’s personal information will be protected, and that the linkage of the information will not result in outcomes against the interest of the individual or their community. If there is an indication that these standards are not being met, public support may erode rapidly, inhibiting or even preventing data linkage research.

Data linkage benefits the population in a number of ways, from identifying trends in disease in a particular population to discovering reasons why children are born with various conditions. The focus of data linkage research must always be for the benefit of the community whilst also ensuring the full protection of an individual’s or community’s identity. There are occasions where complex linkage projects may increase the risk of identities being discovered; however, these risks can be mitigated through rigorous privacy mechanisms.

**FINDING 6**
Robust privacy protection is critical to ensuring the success and survival of any data linkage function.

4.2 Current privacy measures for data linkage

Whilst WA does not have privacy legislation in place, the DLB has a number of mechanisms that ensure the protection of data, which are:

1. Data custodians must be satisfied that the data will remain private when it is released;\(^{21}\)
2. DoH HREC considers any potential risk to privacy in making its determination;\(^ {22}\)

---


3. The DLB ensures that the identifying and clinical parts of data sets can never be accessed together by anyone (the separation principle);\textsuperscript{23} and

4. DLB has in place robust electronic data security protocols.

The rigorous application of these mechanisms has ensured that no known privacy breaches have occurred from data linkage operations in WA.

**FINDING 7**
The current procedures and technical processes applied by the DLB currently provide a high level of privacy protection.

Robust privacy protection is critical to ensuring the success and survival of any data linkage function. Given the importance of the privacy of data used in data linkage, and the success of the current processes, these methods should continue, with two caveats. First, the systems in place should be updated as necessary to counter emerging threats. Second, any duplication occurring at any stage of the process should be streamlined, including around the assessment of privacy as addressed further in Chapter 5 – *Ethics*.

**RECOMMENDATION**
The DLB, and any future linkage facility, continues to apply the current procedures for ensuring data security and privacy, and update them as required. (Recommendation 2)

4.3 **Future privacy measures for data linkage**

Throughout the consultation process for this review, privacy legislation was identified as potentially beneficial to a number of aspects of data linkage. These were said to include:

- greater privacy protection for individuals;
- clearer data use and release criteria;
- greater compatibility with legislated protection of privacy in other jurisdictions (all other Australian States and Territories except WA and SA have privacy legislation); and
- greater public trust in data sharing and data linkage.

\textsuperscript{23} 2016 Data Linkage Review – Submission Process
**Privacy protection**

As noted in 4.2 above, the current privacy measures applied by the DLB are effective. However, this may not be the case with all Government agencies, and particularly those that deal with sensitive data infrequently. Privacy legislation would create a high level framework enabling the legality for releasing data to be assessed, ensuring that standards applied for data release are consistent across Government.

Further, if the legislation was to include penalties for breach of the provisions, as occurs with privacy legislation in other jurisdictions, it would provide a strong disincentive for the malicious access of private data. Currently, there are no general criminal penalties in WA for such breaches.

**Clear data use and release criteria**

Privacy legislation may provide high level guidance as to when data access can be granted; however, specific data access legislation, as proposed in Chapter 6 – *Data Availability and Access*, may provide more detailed guidance around the circumstances in which data can or cannot be released, and for which purposes it may be used.

**Inter-jurisdictional sharing**

Access to a wide variety of data held by the Commonwealth, including Pharmaceutical Benefits Scheme (PBS) and Medicare Benefits Schedule (MBS) data, as well as data held by other States, would be beneficial to users of the DLB. However, submissions to this review have suggested that the lack of privacy legislation in WA is impeding access to Commonwealth and other jurisdictions’ data.

Most privacy legislation in Australian jurisdictions, including the Commonwealth’s *Privacy Act 1988*, require that the jurisdiction receiving data ensures that it be subject to similar, or more stringent, privacy protections than are in place in the donor jurisdiction. While this can be achieved through privacy legislation, Acts typically make provisions for other forms of protection, such as contractual clauses. Therefore, WA’s lack of privacy legislation is not a legal inhibitor to cross-jurisdictional sharing. However, it appears to have caused reluctance by some jurisdictions to share data with WA, which in turn is hampering research initiatives in this State.
It should be noted that the Department of the Prime Minister and Cabinet is strengthening the processes and guidelines around Integration Authority Accreditation to provide assurance that integration of data is undertaken in a safe and secure manner by organisations with requisite skills, processes, infrastructure and culture to undertake high risk projects. As a result of recent changes to the guidelines, Government departments in WA are now eligible (with conditions) to apply for accreditation, which will go some way to improving access.

*Public trust*

As noted in 4.1 above, public trust is essential for maintaining a successful data linkage regime. Privacy legislation that: (i) increases the strength of the privacy regime around public data; and (ii) clearly articulates when it can and cannot be released, will reinforce public trust in the State’s data sharing regime.

**FINDING 8**
The lack of privacy legislation in WA is not a legal inhibitor to data sharing within WA and with other jurisdictions. However, there is a perception externally that the lack of privacy legislation in WA is a concern. There is an opportunity to strengthen the State’s capacity for data linkage through privacy legislation.

With the growing demand for data linkage in WA, the list of Government agencies and other organisations sharing their data will grow. Whilst privacy legislation is not essential in the immediate term, it will be critical to consider its importance for a whole-of-Government approach in the future. This will also assist WA to participate fully in inter-jurisdictional projects and have better access to data.

The assessment of privacy for data linkage requests is addressed in Chapter 5 – *Ethics*.

**RECOMMENDATION**
The Department of the Premier and Cabinet and the Office of the Government Chief Information Officer seek approval from Cabinet to draft privacy legislation for Western Australia. (Recommendation 17)
Privacy Preserving Record Linkage
This enhanced security method for protecting personally identifying information used in linkage is currently being explored by the CDL in collaboration with the DoH. This methodology involves creating encrypted data linkage keys and performing the linkage using the encrypted keys from different datasets, such than no personal identifiers are exchanged in the process and the quality of linkage is maintained. While ‘probabilistic’ linkage using personal identifiers remains the gold standard, the Privacy Preserving Record Linkage would enable the creation of linked datasets from different government agencies in advance of changes to legislation. The current trial involves linking DoH and Medicare data, which are data sets that current governance processes would not otherwise allow to be linked. If these trials are successful, further use of this technique should be explored.
Chapter 5 – Ethics

RECOMMENDATIONS

• The Department of Health Human Research Ethics Committee, and any future Human Research Ethics Committee considering applications for data linkage, be flexible in its scope and operation. For example, it should consider programs of work, as well as discrete projects, and enable ongoing access to linked data sets. (Recommendation 3)

• The Department of Health, and all other State Government agencies, review and amend policies and procedures so that data custodians do not consider privacy and ethical issues (which includes research merit) when approving a data linkage application. This responsibility rests with the Human Research Ethics Committee. (Recommendation 10)

• In conjunction with establishing the data linkage Statutory Body, a Human Research Ethics Committee with specialist expertise in data linkage should be created. The Human Research Ethics Committee should have responsibility for the privacy and ethics approval processes for data linkage, whether or not done for research purposes, and be compliant with the National Statement on Ethical Conduct in Human Research. (Recommendation 23)
5.1 Current ethics review processes

The current data linkage regime within WA involves a HREC which is responsible for ‘reviewing research proposals involving human participants to ensure that they are ethically acceptable and in accordance with relevant standards and guidelines.’ The NSECHR, issued by the NHMRC, is the standard by which Australian HREC’s are registered and certified. The NSECHR sets out the requirements and responsibilities for the establishment of HRECs, the submission of research proposals and the decision-making processes of HRECs. The DoH HREC is registered under the NSECHR.

For the DLB to perform data linkage, the project must first receive ethical approval from the DoH HREC. This must occur regardless of the nature of the project, or any other approvals it has received.

Ethics review committees

It is often the case that projects involving data linkage are required to receive ethics approvals from other NSECHR-registered bodies, such as the WA Aboriginal Health Ethics Committee, prior to receiving approval from the DoH HREC. Some submissions to this review have suggested that the current approvals process could be streamlined if the DoH were to accept the ethics approvals from other bodies, through a mutual recognition arrangement.

It is understood that ethics committees often have specialist knowledge about particular groups or technical processes (e.g. the DoH HREC is one such body having an in-depth knowledge of data linkage). This is highly appropriate given the nuances and complexities associated with many requests for linked data, especially where research in a particular area requires the ethics approval from a specific body that is not the DoH HREC.

However, it may be useful to explore the benefits of mutual recognition ethics processes for national projects, through partnerships between the DoH HREC and other specialist committees in other jurisdictions. This, however, should not overshadow the importance of the current arrangement within WA where circumstances may require more than one ethics approval.

---

FINDING 9
Researchers are sometimes required to seek approvals, for a single project, from a number of different research ethics bodies.

Several submissions to this review expressed concern that the current ethical review processes do not allow for the full and efficient use of linked data. More comprehensive use of linked data would capture the following:

- enhanced use of linked data for the development of broader Government policy;
- programs of work consisting of multiple linked projects; and
- greater access to more regularly updated, and potentially real-time, data sets.

FINDING 10
With greater recognition of the potential of data linkage to serve a variety of purposes beyond traditional research (e.g. reporting, policy and program development) it will be important for ethics approval processes to be designed to allow data linkage capability to realise its full potential.

If WA is to fully exploit its data linkage capabilities, it will be essential to have ethical oversight that provides approvals for projects that may differ greatly from the medical research applications that are currently being evaluated.

To create a shift from the present approach and to cater for the emerging needs of stakeholders, in the short term, a number of changes could be made, including:

1. a single ethics approval for a program of work that may encompass a number of discrete projects, i.e. approval of a program would mean that separate ethics approvals would not be required for each of its related projects; and
2. a single ethics approval for a project, or program, that provides the user with ongoing access to the linked data over an extended period of time.

RECOMMENDATION
The Department of Health Human Research Ethics Committee, and any future Human Research Ethics Committee considering applications for data linkage, be flexible in its scope and operation. For example, it should consider larger programs of work, as well as discrete projects, and enable ongoing access to linked data sets. (Recommendation 3)
Data custodians

Prior to an application for data linkage progressing to the DoH HREC, it must receive in-principle support from all relevant data custodians. This support requires that data custodians ‘undertake a risk assessment to ensure the identification of individuals, patients and Health Service Providers have been considered and appropriately managed prior to releasing of information.’

In addition, the DoH HREC, in giving its approval, must conduct:

“Review of projects requiring the use and disclosure of personal health information without consent to ensure:

- the public interest in the project outweighs the public interest in the protection of privacy;
- the project cannot be conducted using non-identifiable information;
- it is impracticable to seek consent from the people whose information is to be used or disclosed;
- the information requested is the minimum necessary to accomplish the purpose; and
- the project ensures the security of the information.”

Consequently, for any given application all relevant data custodians, along with the DoH HREC, are required to consider the privacy of the data prior to its release. This duplication could give rise to significant delays. There are also concerns that this system ‘undermines the role of the DoH HREC and belies the clinical expertise and experience of the researcher.’

FINDING 11

There is a lack of clarity regarding the roles and responsibilities of the data custodians and the DoH HREC, particularly whether data custodians or the DoH HREC have responsibility for ensuring the privacy of individuals. This duplication is giving rise to delays.

---


28 2016 Data Linkage Review – Submission Process
The responsibility for ensuring the privacy of data should fall to a single body so that the process is streamlined, consistent and clear. For the following reasons, the responsibility for privacy of data should rest with the DoH HREC, or any HREC constituted under future data linkage bodies:

- HRECs have a wide variety of expertise that allows them to consider issues around privacy more widely than data custodians, who do not necessarily receive training or have qualifications related to ethics or privacy;
- HRECs are required to consider the entire set of linked data that will be released to researchers, whereas data custodians are responsible only for their own data sets; and
- HRECs operate as a committee, which ensures wider scrutiny of the decision-making processes, whereas data custodians work alone.

The role of the data custodian should be limited to providing advice to the DoH HREC on issues that concern their specialist knowledge of the data set, and should not include the assessment of privacy.

**RECOMMENDATION**
The Department of Health, and all other State Government agencies, review and amend policies and procedures so that data custodians do not consider privacy and ethical issues (which includes research merit) when approving a data linkage application. This responsibility rests with the Human Research Ethics Committee. (Recommendation 10)

### 5.2 Future ethics review processes

Should the Statutory Body, proposed in Recommendation 22 (Chapter 3 – Governance) be created to take over the responsibility for providing future data linkage services for whole-of-Government, it will be necessary to constitute a HREC. Provisions within the Statutory Body may require that the current DoH HREC sit outside of DoH to cater for a broader-than-health approach.

**FINDING 12**
Should a Statutory Body be established for data linkage (as per Recommendation 22) then a new HREC will need to be created.
This HREC should be compliant with NSECHR, and have specialist expertise in data linkage. In addition, it should adopt the changes that have been proposed to the DoH HREC, namely it should be flexible in its operations.

**RECOMMENDATION**
In conjunction with establishing the data linkage Statutory Body, a Human Research Ethics Committee with specialist expertise in data linkage should be created. The Human Research Ethics Committee should have responsibility for the privacy and ethics approval processes for data linkage and be compliant with the National Statement on Ethical Conduct in Human Research. (Recommendation 23)
Chapter 6 – Data Availability and Access

RECOMMENDATIONS

- Greater awareness of the benefits of sharing data is required across the State Government to improve the availability of data for data linkage. (Recommendation 11)

- The Office of the Government Chief Information Officer develop a set of whole-of-Government standards or guidelines for the collection of data across Government, and facilitate training and support to public sector employees. All State Government agencies should be required to follow these guidelines and will need to make appropriate resources available. (Recommendation 12)

- The Department of Premier and Cabinet and the Office of the Government Chief Information Officer investigate the feasibility of developing policy or legislative mechanisms to better enable the sharing and linking of State Government data. (Recommendation 18)

- The State Government improve its engagement with the not-for-profit sector to enable the sector to participate better in data linkage and analysis. (Recommendation 13)

- State Government agencies are encouraged to explore the feasibility of accessing data collected by third parties e.g. service providers contracted by Government and other private sector organisations. Care would need to be taken if considering outsourcing data collection to the private sector, particularly with respect to ensuring data quality, value for money, and full and continued access to the data by Government. (Recommendation 19)
6.1 Data access and data linkage

A successful data linkage system, particularly a whole-of-Government model, depends on the scope and quality of the data to which it has access. While WA’s data linkage system has largely utilised data collected by the State Government, a variety of other data sources are also available and have the capacity to further increase the complexity of data for research and policy development in the State. These include the data sets of:

- the Commonwealth Government;
- the not-for-profit sector; and
- the private sector.

In addition, access to State Government data, with varying ease of access between agencies and data sets, also remains a challenge that needs to be addressed.

6.2 Access to data held by WA Government agencies

The State Government collects a significant volume of information about WA and its citizens through ongoing service delivery and targeted data collection. The data held by the State is unique, both in the time period covered (some data sets go back as far as the 1960’s) and also in the nature of studies that complement research, such as longitudinal studies like the Raine Cohort Study and the Busselton Health Study.

The benefits of in-depth collections grow exponentially with data linkage and enable the development of a more refined picture of the State. When accessible, these resources are of considerable value to researchers undertaking a wide variety of studies, as well as policy makers working to optimise systems across Government. For more information about WA’s data sets, and their potential to improve the health and wellbeing of its citizens, refer back to Chapter 2 – Background.

Notwithstanding the widespread benefits of linking data within Government agencies, numerous submissions to this review provided examples where gaining access to State Government data was exceptionally difficult. There are a wide variety of reasons for this, some of which are addressed in Chapter 2 – Privacy and Chapter 3 – Ethics.
This Chapter will focus on issues that arise across Government and do not appear to be confined to particular agencies. These are:

- variability in data collection and management across the public sector;
- data quality issues;
- concerns around the misuse or misinterpretation of data; and
- lack of clarity around the circumstances in which data can be released.

**Variability in data collection, management and quality, and concerns around misuse of data**

There is a concern among public sector agencies around the potential misuse or misinterpretation of Government data. This issue was raised in a submission to this review and also via anecdotal evidence received by the Advisory Group.

The Advisory Group does not consider that such concerns justify a refusal to release data for two reasons. Firstly, this argument could be used to protect an agency from normal and appropriate review of its practices. Secondly, one way to determine if the analytics applied to the data does amount to a misinterpretation is to expose it to further examination. However, it is entirely appropriate for agencies to have a period of privileged access to reports prepared by analysis of datasets that are not in the public domain. It would be expected that any comments agencies make on such reports are considered by the author, and justifiable changes are made before public release.

**FINDING 13**

There are concerns within many State Government agencies, particularly those exposed to media scrutiny, that data will be misinterpreted if released.

A number of submissions to this review also identified the failure of some State Government agencies to view data management as ‘core business’ and the negative effect this has on the use, sharing and linking of data. These submissions outlined the difficulties experienced in accessing data, and problems with the quality of data collected and maintained by agencies.

For example, Landgate has experienced difficulties with the quality and format of the data provided by Government agencies for its SPUR program, making it difficult to use. Furthermore, some data collecting agencies acknowledged that competition for resources within their agency has meant that investment in managing their data was often sub-optimal.
The value of data is largely contingent on its quality. Poor quality data is likely to be misleading when analysed and often cannot be accurately linked. A number of submissions to this review identified concerns around the quality of some data held across a range of Government agencies. Given the variety of submissions received on data quality, this appears to be a sector-wide issue, and not restricted to any particular agency. The failure to ensure data collection is consistently performed at a high quality harms the chance of deriving tangible benefits.

<table>
<thead>
<tr>
<th>FINDING 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is variability in data collection and management practices across the public sector, with some agencies not viewing this function as a priority nor allocating adequate resources. This affects the consistency, availability and quality of data for data linkage.</td>
</tr>
</tbody>
</table>

The State Government is aware of these issues and has taken some steps to mitigate them. The WA Government Chief Information Officer (GCIO) was appointed and a supporting office, the OGCIO, was established as part of the Government’s initiative to improve the capabilities necessary to generate a sophisticated and efficient approach to utilising data across the public sector. Some of the current OGCIO key initiatives are:

- Whole-of-Government Open Data Policy (completed on 3 July 2015): Aimed at improving the management and use of public sector data assets, and including greater release of appropriate data. The OGCIO is working together with Landgate to implement this policy.

- Whole-of-Government Data Classification Policy (due 2016): This policy is expected to provide a framework on the treatment of Government data for the purpose of information sharing with Government entities, suitable organisations and the community. Agencies will benefit from improved data sources to strengthen planning and decision-making, which in turn will support the delivery of better products and services.

- Other policy areas currently under development include the Privacy Policy (due in 2017) and the ICT Procurement Policy (due in 2017).

---

Key policy areas that are already complete and available on the OGCIO website include Open Data, Digital Services, Digital Security, Cloud, Interoperability and Business Continuity and Data Recovery.

These measures represent a positive start. However, these actions alone do not fully address the inconsistency and lack of priority given to data management by some Government agencies.

There are two ways that the Advisory Group considers the quality of data across Government can be improved. The first is in the increase of open data. Work done by both the OGCIO and Landgate in this area demonstrates that opening data for broader use can lead to better quality data. The second way is to standardise data collection across Government. Data is currently collected in varying shapes, sizes and quality, creating scenarios where linkage can be difficult, costly or impracticable. Data idiosyncrasies are the reason data linkage is not always viable; hence, the more work done initially to standardise data collection, the better the chances of successful linkage. The Advisory Group is of the opinion that a set of standards, or guidelines, should be developed to guide agencies involved in the collection of data sets. These guidelines should provide a common platform for data to be collected State-wide, thus allowing for the early detection and correction of errors commonly committed during the collection stage and increasing the likelihood of successful linkage in the future.

**RECOMMENDATION**
Greater awareness of the benefits of sharing data is required across the State Government to improve the availability of data for data linkage. (Recommendation 11)

**RECOMMENDATION**
The Office of the Government Chief Information Officer develop a set of whole-of-Government standards or guidelines for the collection of data across Government, and facilitate training and support to public sector employees. All State Government agencies should be required to follow these guidelines and will need to make appropriate resources available. (Recommendation 12)

*Policy and legislative regimes around access to data*
One of the most significant difficulties preventing data sharing and linkage activities is the dense legislative framework and the policy uncertainty around the circumstances for data release. As the instruments vary widely between agencies and data sets, there is no single release standard.
Another complicating factor is the number of administrative instruments that must be considered in deciding to release data. For example, in the DoH, nine criteria were cited which data custodians have to consider before releasing their data.  

**FINDING 15**

There are divergent legislative and policy regimes between agencies in relation to data access. This creates inconsistency across the sector and, in some cases, inhibits data linkage and data use.

The problems associated with data linkage are varied and multifaceted. While a variety of solutions have been proposed in this review, the Advisory Group is of the opinion that none of the recommendations alone will be sufficient to release data held in some of the less willing agencies. As a result, it is recommended for State Government to study the possibility of passing legislation that facilitates access to data.

Other jurisdictions have implemented this with some success. In 2015, NSW enacted legislation facilitating access to Government held data, through its *Data Sharing (Government Sector) Act 2015 (NSW)*; however, it is considered to be fairly limited in scope. In that instance, access to data can only be obtained where the Premier has advised the relevant Minister that the data concerned is required to be shared for the purpose of advancing Government policy, and can only be used in the State’s Data Analytics Centre. While anecdotal evidence from NSW suggests that this has assisted in creating a more open data culture in agencies, the powers are still considered to be very narrow.

The multifaceted problems associated with data access in WA will require a broader approach than that adopted in NSW. There is a need for WA to move beyond agencies acting as the gate keepers to “their” data, to a position where the release of information is approved subject to clearly specified criteria being met. The most appropriate mechanism to achieve this is through legislation and establishing the criteria under which data held by Government can be accessed.

---

30 *2016 Data Linkage Review – Submission Process*
Such legislation should seek to:

- clearly articulate the circumstances in which Government agencies can release data;
- provide a right of appeal to an appropriate body if release of data is refused;
- make provisions for access to data within a timeframe on the request of certain individuals; and
- take into consideration relevant privacy legislation to ensure that no individual’s privacy is breached or put at risk as a result of data releases.

It is also noted that in South Australia, the Public Sector (Data Sharing) Bill 2016 was introduced in Parliament in August 2016 to facilitate the sharing of data between public sector agencies, and public sector agencies and other entities, and to provide for an Office of Data Analytics.

**RECOMMENDATION**
The Department of Premier and Cabinet and Office of the Government Chief Information Officer seek approval to investigate the feasibility of developing policy or legislative mechanisms to better enable the sharing and linking of State Government data. (Recommendation 18)

### 6.3 Access to data collected by the Commonwealth Government

Valuable, and often complementary, data is also collected by the Commonwealth Government, both in the medical sphere (through data sets such as those for the PBS and the MBS) and across a range of other areas.

In the past the Commonwealth has demonstrated a reluctance to share data and engage in data linkage with other jurisdictions, particularly WA. The justification given is mainly due to the lack of privacy legislation (covered in more detail in Chapter 4 – *Privacy*). However, this may not be the only reason, given the difficulties other jurisdictions with privacy legislation have also had in accessing data.  

---

31 2016 Data Linkage Review – Submission Process
Recent changes have been made to strengthen the process and criteria for Integrating Authority Accreditation. This may lead to greater access to Commonwealth Government data sets for WA Government agencies that were previously ineligible to apply.

The Advisory Group also notes the work that is currently occurring through the Council of Australian Governments (COAG) process, which substantiates the Commonwealth’s willingness to engage with the States and Territories about sharing data.

6.4 Access to data held external to Government

Data held outside Government is valuable as it provides unique insights into how the services provided by these sectors link with the work of Government. Two other key collectors and holders of personal data are the not-for-profit and private sectors.

The nature of the data collected by these organisations, the outcomes they are seeking to achieve, and their relationship to Government, vary significantly. For these reasons they must be considered separately. It is possible that these sectors could use data linkage to further refine their services or develop discrete products.

There are two broad categories into which this data falls: 1) data collected by not-for-profit organisations that are seeking to link and use their data to better align their services with Government services; and 2) data collected by private organisations through either the normal operation of their business or for the explicit purpose of sale.

Data collected by the not-for-profit sector

Not-for-profit organisations provide a valuable source of information about members of the community who are considered “at risk”. These groups would also benefit greatly from access to linked data services, including the ability to delivering more effective and better targeted services to the community.

Submissions from not-for-profit groups to this review indicated obstacles in engaging with data linkage and the DLB. Time and cost were identified as issues which are exacerbated for not-for-profit organisations. A lack of expertise and experience in engaging with the DLB also posed a problem, as well underdeveloped capabilities and capacities across the sector.
FINDING 16
A significant amount of data is collected by the not-for-profit sector; however, these organisations may not have the capacity to manage data in a way that can be readily linked.

Improving engagement with the not-for-profit sector could be advanced in a number of ways. This includes increasing the speed and decreasing the cost of data linkage to make it more accessible. The second would be to initiate an active program of engagement with the not-for-profit sector to communicate the benefits of data linkage and encourage community service organisations to link their data into WADLS.

The WA Government’s Partnership Forum, a body aimed at improving the relationship and cooperation between the public and not-for-profit community service sectors, has identified data sharing and data linkage as key areas to improve analysis and community services’ efficiency and effectiveness in its Strategic Directions for 2016 and 2017.

RECOMMENDATION
The State Government improve its engagement with the not-for-profit sector to enable the sector to participate better in data linkage and analysis. (Recommendation 13)

Data collected by the private sector
As the value of data becomes more widely recognised by Government, and the amount of data collected by the private sector continues to grow, it is anticipated there will be significant increases in the degree to which data is traded. For example, a submission to this review raised the possibility of purchasing data to perform data linkage in a spatial context.

This may represent an opportunity for Government to access and use a broader range of data sets than it currently holds and to reduce some agencies’ data collection activities (and the associated costs) where this can be done more efficiently by non-Government organisations.

FINDING 17
There are opportunities for the State Government to save on data collection and expand the data sets available for use where similar or complementary data can be purchased or obtained from private sector organisations (for example, private health insurance providers or Google Maps).
Government, researchers and others stand to gain significantly from more cost-effective methods of data sharing between sectors. However, care would need to be taken when considering outsourcing to the private sector. This includes thorough consideration of the actual net benefit of outsourcing the data collection, and the rights and assurances required with respect to full and continuing accessibility of the data to Government and data quality.

Some of these issues can be improved through comprehensive contractual arrangements with protections addressing the above, however it may not solve all and therefore the decision to outsource must be made carefully.

**RECOMMENDATION**
State Government agencies are encouraged to explore the feasibility of accessing data collected by third parties e.g. service providers contracted by Government and other private sector organisations. Care would need to be taken if considering outsourcing data collection to the private sector, particularly with respect to ensuring data quality, value for money, and full and continued access to the data by Government. (Recommendation 19)
RECOMMENDATIONS

- The Data Linkage Branch, in conjunction with key stakeholders, develop and publish key milestones, and their corresponding anticipated timeframes, for delivering data linkage services. (Recommendation 4)

- The Data Linkage Branch develop an automated online linkage project application and tracking system. Up-to-date information on application status, milestones, actions required and anticipated completion dates should be included in this system and be fully available to applicants. (Recommendation 5)

- The Data Linkage Branch continue delivering client services to applicants and stakeholders, including providing feedback on draft applications for data linkage services. (Recommendation 6)

- The Data Linkage Branch provide additional guidance to clients, particularly those new to the service, through training and more fully annotated application documents. Similarly, Government and research organisations implement procedures to ensure appropriate quality of applications for linked data. (Recommendation 7)

- The Department of Health, and all other State Government agencies, review and amend policies and procedures so that data custodians do not consider privacy and ethical issues (which includes research merit) when approving a data linkage application. This responsibility rests with the Human Research Ethics Committee. (Recommendation 10)

- The Data Linkage Branch, through the Department of Health, and in conjunction with other State Government agencies, invest in increasing the number of data sets that use the Custodian Administered Research Extract Server. (Recommendation 15)

- The Data Linkage Branch, in conjunction with key stakeholders, pilot the use of data repositories for commonly linked and used data sets. The potential for greater use of data repositories should be evaluated as part of this project. (Recommendation 16)
7.1 Introduction

A good process should always focus on delivering value to its clients/users. Good processes should never sit in isolation and should always flow smoothly from beginning to end, through clear definition and sequencing of tasks, to ensure the ideal resources and environment are in place to achieve its objectives.

Users of the data linkage service, administered by the DLB, expect a ‘linked data’ product/service that is relevant to their research needs, is cost-effective and delivered in a timely manner.

However, users have raised concerns about the long timeframes associated with accessing linked data. A number of submissions to this review noted cases in which it could take applicants several years to receive data, including for a project that DoH funded.

Significant differences in timeframes were also noted between linkage services provided by the DLB and linkage services delivered in other jurisdictions. For example, a submission to this review highlighted a case where:

“The time between initial submission of the EOI to full data custodian approval in WA took 22 months. In comparison, the same request from the Centre for Health Record Linkage (CHeReL) for counterpart data from NSW took <2 months”

Time delays present difficulties for researchers attempting to secure grant funding for their projects. Time delays also reduce their capacity to publish in a timeframe that makes policy making useful. These difficulties have started to erode the State’s research competitiveness and reputation, along with its ability to deliver benefits to the community.

In addition to time delays, a number of issues have been raised as to the appropriateness of some decisions being made, which include: a) data custodians potentially making judgements outside their scope of responsibilities; b) a convoluted and unclear approvals process; and c) duplication throughout a number of stages of the data linkage process. This all equates to increased time delays along the process but also gives rise to client/user frustration as they have to deal with a lack of clarity, openness and structure.

32 2016 Data Linkage Review – Submission Process
To ensure an efficient and effective data linkage process it will be critical for these issues to be addressed.

The Advisory Group notes that the DoH and DLB have acknowledged concerns around delays, particularly around more difficult cases, and have conducted an internal review to address this – a welcome development that, together with this review, is expected to improve timely access to linked data.

7.2 The data linkage process

The overview of the DoH data linkage process, as shown in Figure 4 below, describes the key steps, including application, approvals, linkage, extraction and quality assurance.

<table>
<thead>
<tr>
<th>Application</th>
<th>Approvals</th>
<th>Part 1 Linkage</th>
<th>Extraction</th>
<th>Part 2 Linkage &amp; QA</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLBL, Researcher and Data Custodian/s work together to develop the draft application</td>
<td>DoH HREC approval</td>
<td>DLB gathers raw data from Data Custodian/s, creates linkage keys then sends to Data Custodian/s</td>
<td>Data Custodian/s attach content data to linkage keys and sends to DLB</td>
<td>DLB merges linkage keys to create the linked data set, quality assures and encrypts</td>
</tr>
<tr>
<td>Data Custodian/s provide in-principle support</td>
<td>Data Custodian/s approval</td>
<td></td>
<td>DLB provides fully linked data set/s to Researcher</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Overview of the DoH data linkage process

The following roles and responsibilities form part of the data linkage process:

**Applicant:** A person or group seeking linked data, also referred to as the client or user.

**DLB:** Administers the data linkage process, works with the applicant providing guidance and support on the development of a draft application. Ensures compliance against DoH policy and relevant legislative frameworks. Progresses the draft application to the DoH HREC and the Data Custodian/s. Extracts linkage keys, performs data linkage and quality checks. Ensures final release of linked data to the applicant.
Data Custodian: The person within an organisation or agency formally assigned to collect, manage, secure and disclose a data set on a day-to-day basis at the direction of the data steward. Works closely with the DLB and the Researcher on the draft application. Provides in-principle support to the draft application and provides final approval of the application. Extracts linkage keys for the purposes of data linkage (where that data is not contained within the CARES).

DoH HREC: A human research ethics body that evaluates all data linkage application based on ethics and privacy requirements. Provides HREC approval to the application.

7.3 Timeframes

Across the five stages of the data linkage process there are a number of critical steps required to ensure the appropriate release of linked data for research purposes.

The process involves a number of stakeholders (as outlined above), all who play a critical role in data linkage, from providing advice and guidance to the researcher to improve the quality of their application, to providing ethics approval and formal approval to release the data, and many stages in between.

Submissions received as part of this review have indicated that many users of the data linkage service are experiencing long waiting times, that is, from the time at which they submit their draft application to the time at which that data is released. These waiting times are highly variable and in many cases excessive.

It has been suggested that the multiple approvals required on some applications is unnecessary and is exacerbating the long waiting times. 33

Applicants have identified two difficulties that arise with respect to the current timeframes. The first is their length. These vary significantly, and while most applications are completed in a reasonable time, some are subject to excessive delays.

---

33 2016 Data Linkage Review – Submission Process
These delays arise at a number of stages in the application process, and consequently are addressed throughout this report. The second is the uncertainty associated with the timelines.

**FINDING 18**
Users of the data linkage service are experiencing long waiting times to receive their linked data, and these waiting times seem highly variable.

**FINDING 19**
There is a lack of transparency regarding the time it will take for the application to progress through all key stages of the data linkage process.

Submissions to this review have also indicated the lack of transparency with regards to the time it will take for their application to be progressed across the five stages of the data linkage process.

**RECOMMENDATION**
The DLB, in conjunction with key stakeholders, develop and publish key milestones, and their corresponding anticipated timeframes, for delivering data linkage services. (Recommendation 4)

### 7.4 Step 1 – Application Stage

As depicted in Figure 4 above, the data linkage process requires all researchers seeking data linkage to submit a draft application to the DLB. The DLB then forwards the draft application to the data custodian who reviews and provides feedback on the application via the DLB. The data custodian at this point also provides in-principle support for the application. The data custodian, the DLB and the researcher work together to refine the application, sometimes making various amendments along the way.

Once in-principle support is provided by the data custodian the formal application is then forwarded to the DLB, which then undertakes a further review before submitting to the DoH HREC.

If ethics approval is provided by the DoH HREC, final approval is then formally sought from the data custodians who then allow release of the data.
If ethics approval is not provided, then the researcher can withdraw the application or work with the DLB to amend the application, and re-submit to the DoH HREC for further consideration.

FINDING 20
The data linkage process operates as a paper-based system which is slow, duplicative and inhibits timely feedback to applicants.

RECOMMENDATION
The Data Linkage Branch develop an automated online linkage project application and tracking system. Up-to-date information on application status, milestones, actions required and anticipated completion dates should be included in this system and be fully available to applicants. (Recommendation 5)

Through the submissions to this review it is clear that the role of the DLB Client Services Team is well received by users of the data linkage service. The DLB takes much time and effort to work with the researchers and the data custodians to ensure that the application is developed and progressed in a timely manner. However, there have been occasions where the quality of the initial application is lacking, and therefore further work is required to improve the application so that it is ready for approval. This additional effort is contributing to time delays across the data linkage process, and the DLB Client Services Team is spending a disproportionate of time on single applications.

FINDING 21
The DLB Client Services Team is adding value to the data linkage process by ensuring applications are of a high quality before they progress for approval, thereby seeking to improve the timeliness of the approvals process.

FINDING 22
Due to the poor quality of some applications there is a disproportionate amount of time spent by the DLB working with researchers and data custodian/s to ensure the development of an appropriate application.

RECOMMENDATION
The Data Linkage Branch continue delivering client services to applicants and stakeholders, including providing feedback on draft applications for data linkage service. (Recommendation 6)
RECOMMENDATION

The Data Linkage Branch provide additional guidance to clients, particularly those new to the service, through training and more fully annotated application documents. Similarly, Government and research organisations, implement procedures to ensure appropriate quality of applications in linked data. (Recommendation 7)

7.5 Step 2 – Approvals Stage

During the application stage, data custodians are required to provide in-principle support to the draft application before it is progressed to the DoH HREC for approval. Data custodians operate under policy, legislative and other requirements, which may differ between agencies and even between datasets within an agency.

Generally, data custodians consider the following:

- the purpose of the project and what data is needed, including whether the data is appropriate for undertaking the project and answering the research questions;
- any risks to the identifiability of individuals;
- adequate data security and data integrity/quality;
- existence of appropriate data management practices, including around retention and disposal of data;
- impact on the agency’s operations; and
- that relevant legislative frameworks and policies are complied with.

Once in-principle support is provided, the DLB progresses the application to the DoH HREC who assesses the application against its governance and ethics framework. Ethics review is considered in Chapter 5 – Ethics. Once DoH HREC approval is provided, the application is then progressed to the data custodian/s for formal approval.

A number of submissions to this review indicated that when data custodians are reviewing applications they are also including an assessment of the research project’s merit (research merit). Submissions also reported discomfort with the lack of transparency around the approvals process, in particular decision-making by data custodians, and the complex nature of the application and approvals systems, with calls for a more streamlined and transparent process.

In practice, research merit is assessed as a component of ethics. It is the role of the DoH HREC to assess the application based on ethics and privacy.
Therefore, to ensure no duplication of the approvals process, and further time delays, the determination of research merit, as part of an ethics assessment, should remain solely with the DoH HREC. This issue is addressed in Chapter 5 – *Ethics* – Recommendation 10. Clear and transparent communication of this process should also be pursued to avoid confusion and allow data linkage clients to better understand the process.

**FINDING 23**
There is a lack of clarity as to who has the responsibility for determining the research merit of an application, whether it be the data custodian/s or the DoH HREC.

**RECOMMENDATION**
See Recommendation 10, Chapter 5 – *Ethics*.

### 7.6 Step 3 – Part 1 Linkage

In this stage the DLB gathers raw data from the data custodian/s to create a linkage key. The raw data usually includes information such as name, date of birth, sex, post code, etc. This data is identifiable; however, it does not include service information, for e.g. relating to a particular health diagnosis, criminal conviction or hospital visit. From this raw data, the DLB creates the linkage key in the form of an ID number. The DLB then forwards the ID number to the data custodian/s. This linkage process is the first of two steps to ensure privacy of the individual's data is protected through a process of de-identification.

See example below:

**Raw Data Provided by the Data Custodian**
- Name: Joe Smith
- DOB: 3/4/1972
- Sex: Male
- Postcode: 1234

**Linkage Key Provided by the DLB**
- Local ID: 1432 (which will point to the raw data above within the DLB system)
7.7  Step 4 – Extraction Stage

At the extraction stage, the data custodian/s attach content data to the linkage key. The content data includes information such as a specific health diagnosis, a criminal conviction, a hospital admission and an emergency department admission. This forms the second step in ensuring data is de-identified and privacy of the individual’s information is protected.

If the research project requires data from multiple sources, which is usually the case, then data from more than just one agency will need to be extracted. Often this stage can take a long time as there are multiple people involved with multiple systems being accessed to generate the data requested. Sometimes the process of extracting the data is one of many competing tasks for the data custodian/s, which makes it difficult for the DLB to determine how long the extraction process will take, and in turn how long the delivery of linked data to the researcher will take.

Over time, the development of the CARES which, through a central hub of data collections, has greatly reduced the time taken to extract data, compared to the current process of data extraction. As a DLB initiative, CARES has been well received by many users of the data linkage system. It also presents as an opportunity to build on this capability to store more Government data collections, so there is less reliance on the current data linkage extraction process.

**FINDING 24**
CARES has been well received by users as it provides a faster, more accurate and predictable approach to linked data extraction.

**RECOMMENDATION**
The Data Linkage Branch, through the Department of Health, and in conjunction with other State Government agencies, invest in increasing the number of data sets that use the Custodian Administered Research Extract Server. (Recommendation 15)

7.8  Step 5 – Part 2 – Linkage & Quality Assurance

As explained above, once the data custodian/s have attached relevant content data to the linkage key, this data is then forwarded to the DLB. Note: as mentioned previously this usually occurs multiple times by multiple agencies depending on the nature of the data requested. Once the DLB receives the linkage keys with content data attached, the information is merged to form the linked data set.
The data linkage set is then encrypted and quality checked to ensure it provides all the information requested in the original application.

The DLB then releases the linked data set to the applicant.

### 7.9 Repository Model for Data Linkage

Although WA’s data linkage system served its original purpose well, the classical data extraction and delivery methods used are not producing outcomes in a timely and cost-effective manner.

Under the current system, in which data is extracted and linked on a project by project basis (and then ultimately destroyed upon completion of each project) data custodians are required to extract the same information on multiple occasions. This places a considerable resource burden on data custodians, resulting in significant duplication of effort and large opportunity costs.

Submissions to this review have recommended investigating the possible creation of enduring linked databases of de-identified information, known as data repositories, where data would be securely stored, extracted and delivered. Under this model, data custodians would only need to extract commonly used data sets once, with users being able to access the same data on multiple occasions (subject to relevant approvals). This system would eliminate the need to recreate commonly-used data and linkage keys, and is expected to reduce the time and costs needed for data extraction and analysis.

The CARES model is an example of the usefulness of streamlined data linkage processes. CARES was created to address increasing scope, number and complexity of DLB’s project requests and has successfully reduced unnecessary duplication of extraction by data custodians. However, under CARES, data is still extracted on a project-by-project basis. Adding repositories of commonly-used data sets to the model would further reduce the burden placed on data custodians and the DLB, in turn decreasing duplication and improving efficiency.

While such an integrated system could reduce inefficiencies associated with the classical model of data delivery, it could encounter difficulties within the policy and political framework affecting data linkage within WA. Current privacy and confidentiality guidelines, as well as requirements regarding separation of identifiers from content, may present barriers for such an integrated system and will require further investigation.
Major repositories of linked data are in practice in both Australia and overseas. For example, New Zealand’s IDI model links longitudinal data sets that cover an extended range of pathways and transitions. The database consists of anonymised links of administrative data from a range of organisations, covering a wide range of fields.

Another example is MADIP, a collaborative cross-portfolio partnership between the ABS and its partner agencies, the ATO and the Commonwealth Departments of Social Services, Health and Human Services. MADIP allows the ABS and its partners to conduct cross-Governmental analysis and evaluation in a secure manner via sharing and integrating sensitive data such as existing surveys, administrative collections and censuses.

Moreover, several submissions to this review suggested consideration of a data lab and/or secure remote access facilities as a means to access underlying granular data while significantly managing the risk of potential re-identification. Additional benefits could include faster access to data and reduced costs and resources when compared to a case-by-case approval and linkage model e.g. if access is granted to accredited researchers and analysts rather than projects. There are further options such as localised ‘data islands’ with no remote access enabled where users would need to be on-site to access information. The benefits and feasibility of data labs or data islands for WA could be further explored in conjunction with repositories.

**FINDING 25**
Repositories of linked data, rather than the use of data linkage keys, are an emerging model of working with linked data that has delivered positive outcomes, in terms of providing ready access to data, in other jurisdictions.

**RECOMMENDATION**
The Data Linkage Branch, in conjunction with key stakeholders, pilot the use of data repositories for commonly linked and used data sets. The potential for greater use of data repositories should be evaluated as part of this project. (Recommendation 16)
Chapter 8 – Funding and Linkage Charges

RECOMMENDATIONS

- The Data Linkage Branch be more transparent regarding its charging for services, including: making its charging schedule and algorithms publicly available; increasing the level of detail provided in invoices; and providing justification to clients where the final charge is more than 10% above the final quote. The Data Linkage Branch should also review its charging schedule to ensure its percentage of cost recovery is consistent for all project types. (Recommendation 8)

- The State Government increase funding to the Data Linkage Branch. Greater resourcing is essential to support and improve data linkage service provision. (Recommendation 14)
8.1 Funding for the DLB and data linkage

In 2015-16 the DLB’s total operating budget was $2.257 million, with expenditure totalling $2.297 million. State Government funding through the DoH was reportedly $1.2 million.34

Over the past five years, DLB funding has originated from three main contributors:

- approximately 50% from recurrent State Government funding;
- approximately 40% from grants received from a variety of organisations; and
- between 5% to 10% from cost recovery by charging for data linkage products and services (it should be noted that as a result of a new pricing policy this will increase to 20-30% in 2016-17).

These three funding sources have sustained the DLB and the WADLS since its inception and have enabled it to deliver linked data products and services to users.

Increasing demands on the DLB, including growth of the system, and the complexity of applications for linked data, have not been matched by increased grant funding or substantial increases in recurrent funding from the DoH, particularly over the last ten years.

The DLB introduced a new pricing policy in November 2015, increasing the level of cost recovery for the first time (from a range of 5-10% to 20-30%). Anecdotal evidence from submissions to this review indicated that the cost of data linkage in WA was already prohibitive and potentially stopping a number of projects from proceeding. It is concluded therefore that an increased cost recovery schedule has the potential to prevent a greater number of research projects from being undertaken.

DoH funding

The DoH provides recurrent funding to the DLB of about $1.2 million per annum, or 50% of the DLB’s total annual budget. The recurrent funding to the DLB provides for ten staff, some equipment and the infrastructure linkages of nineteen data sets.35 Figure 5 below shows the funding allocated to the DLB from 2007/08 to 2015/16.

---

34 2016 Data Linkage Review – Submission Process
35 See note 34
Figure 5: DLB funding between 2007/08 and 2015/16 – Note that in kind contributions are not included in the funding table above.36

Grant funding

Over the last few years grant funding made up approximately 40% of DLB’s total annual budget. Grant funding has been provided from a range of sources, including:

- DPP;
- PHRN;
- Road Safety Commission (funding provided for a program of work related to Road Safety since 2007); and
- Lotterywest (in partnership with Curtin University, UWA and TKI).

The DLB indicates that while grant funding is of assistance to its activities, it is often time limited and allocated for a specific purpose37.

36 2016 Data Linkage Review – Submission Process
37 See note 36
To the surprise of the Advisory Group, further grant funding was offered to the DLB but was ultimately not taken up. Specifically, the DLB did not access funding offered through PHRN due to difficulties in negotiating a mechanism for funding provision and DoH concerns over the terms of the Participation Agreement. DoH’s areas of concern included consistency between agreements, ownership of intellectual property, provision of information to third parties and reporting requirements. All other jurisdictions in Australia were able to accommodate the requirements of the PHRN and received additional funding.

The Advisory Group notes that whilst grant funding is a very useful supplement, it is not a replacement for sustainable long term funding for core activities and strategic growth. However, the Advisory Group would encourage that, wherever possible, grant funding be actively accessed and that the issues pertaining to such funding be resolved as a priority.

8.2 Cost recovery through charging for data linkage products and services

The DLB has always operated on a partial cost recovery basis. The introduction of the new pricing policy (Charging and Access Policy) in November 2015 was expected to increase the amount recovered from linkage activities from 5-10% of the DLB’s annual budget per year to 20-30% per year. This will significantly increase the price of linkage services.

The new pricing schedule was flagged as a potential issue in submissions to this review. The Advisory Group also notes that the new charging schedule could place a disproportionate burden on some project types, which could risk a number of projects not going ahead due to the cost.

For example, in a submission to this review, a Government agency indicated that an application made under the DLB’s new Charging and Access Policy was quoted at $79,652. It is understood that a number of alternative pricing options were presented by the DLB, however choosing a lower cost option meant a reduction in scope and for this reason the agency chose not to proceed.

---

38 2016 Data Linkage Review – Submission Process
The Charging and Access Policy was accompanied by updates to the charging schedule used by the DLB. The policy resulted in a number of changes to the data linkage regime, specifying that charges should be applied to all users, including those within the DoH, and included charges for all key DLB products and services. While the Charging and Access Policy is in the public domain, the charging schedule is not.

Anecdotal evidence from the submissions to this review suggested that the price of performing linkage under the new schedule is approximately two to three times more than it was under the old pricing structure. This is in line with the DLB’s modelling. This is a significant increase in charges and will inevitably have an impact. The Advisory Group notes that the price of data linkage services is largely dependent on the degree to which the State, or the DoH, subsidises the service.

It is further noted that researchers are able to offset these costs with grant funding and agency agreements, provided they are aware of the total cost figure. A combination of partial cost recovery and Government subsidy seems to be the best approach to tackle these difficulties, and is in line with policies in other jurisdictions.

**FINDING 26**

Under the new pricing schedule for data linkage services, there have been significant increases to the charges levied by the DLB for some linkage activities.

In addition to the general increase in price, the policy places a disproportionate burden on some project types. For example, the pricing schedule for ad-hoc linkages is broken into three bands based on the number of records to be linked: small linkage, medium linkage and large linkage. All use the logarithm of the number of records to be linked, multiplied by a value based on the band in which it falls. The multiplier increases between each band, however at a decreasing increment. But, as the multipliers are substantially different from one another, there is the potential for major increases in cost for marginal increases in the complexity of work e.g. for projects near a band threshold.

**FINDING 27**

The DLB’s current charging schedule may not be proportionate to the work performed for certain types of linkage, placing undue burden on some projects.

---

39 2016 Data Linkage Review – Submission Process
As noted in Chapter 3 – Governance, the transparency of actions taken within the DLB was raised as an area of concern in a number of submissions to this review, particularly about the amount of information provided regarding pricing structures and charges used by the DLB for its services.

The DLB has not made its charging structure publicly available. While it does provide the Access and Charging Policy, this does not set out a schedule of fees for its services, rather it provides general principles to guide the way charges should be set. As a result, it is difficult for applicants to predict or budget for the total cost of the linkage, or explore ways to minimise the cost. The issue of transparency is exacerbated by a lack of reliability around quotes provided by the DLB, including the potential for an estimate to change due to the time it takes to receive data, and the minimal information provided in quotes or invoices. Transparency enables applicants to better understand and even scrutinise the charges applied. Openness creates an opportunity for feedback from clients and stakeholders which could give rise to further optimising the more resource intensive areas of the DLB’s operations and the charges applied.

**FINDING 28**
The level of openness and transparency currently offered by the DLB in charges for data linkage is causing dissatisfaction among clients of the service.

**RECOMMENDATION**
The DLB be more transparent regarding its charging for services, including: making its charging schedule and algorithms publicly available; increasing the level of detail provided in invoices; and providing justification to clients where the final charge is more than 10% above the final quote. The DLB should also review its charging schedule to ensure its percentage of cost recovery is consistent for all project types. (Recommendation 8)

---


41 2016 Data Linkage Review – Submission Process
8.3 Funding and charging in other jurisdictions

The most recent national funding data found by the Advisory Group is shown in Figure 6 below.

![Figure 6: Funding received by State data linkage units between 2011 and 2013](image)

Note the funding component illustrated above includes grants and cost recovery (if applicable). Data does not include figures for national data linkage units.

Direct funding ranged from $0.7 million for Tasmania to around $6.15 million for NSW, averaging at $3.5 million per State. With the exception of WA and SA, “in-kind” funding was significantly lower than direct funding, averaging at $1.8 million per State. There are also many factors to be taken into account when examining Figure 2, such as WA and NSW having an established versus an emerging linkage unit, the population sizes of the different jurisdictions, the scale of linkage undertaken etc.

---

42 Tew, M, Dalziel, KM, Petrie, DJ & Clarke, PM, 2016, ‘Growth of linked hospital data use in Australia, a systematic review’ Australian Health Review, p. 4
Interestingly, the two States who are responsible for the majority of research publication output (WA and NSW) have quite different funding structures, with WA’s ratio of funding to in-kind being 50% whereas for NSW/ACT it is 15%. WA’s high ratio of “in-kind” funding may be an obstacle to future investment in the expansion of WA’s data linkage capabilities.

When evaluating the charges for data linkage, comparing WA to other jurisdictions is difficult. Some jurisdictions have linkage systems that are much less complex and refined than those in the DLB, and some may not provide the same level of quality linkages offered by the DLB (e.g. less administrative checking). All health related data linkage units in Australia receive a degree of Government funding for their operations. Nevertheless, the charges for data linkage have significant bearing on the scope, volume and quality of research that is conducted in the State, and unless competitive pricing policies are set, WA will not attract the highest calibre research.

The most comparable Australian linkage unit to the DLB is NSW’s CHeReL. Although specific information on the pricing policy applied by CHeReL is not available, the Advisory Group has received a number of anecdotal comparisons of the costs to performing linkage in the two jurisdictions. With the exception of one example, all have indicated that the cost of performing linkage in WA is more than in NSW for a comparable project (approximately two to three times that of NSW). Stakeholders have also indicated that this price differential (combined with the dramatic difference in time to obtain data between NSW and WA) has caused some researchers to redirect their linkage applications away from WA to NSW.

### 8.4 Funding and charging for data linkage in WA moving forward

For the DLB to continue to both deliver its services in a cost effective manner and upgrade its systems, there must be greater investment in the form of secured ongoing funding.

The degree to which this funding will need to be increased will depend upon circumstances that the Advisory Group does not have the ability to predict, such as the DLB’s capacity to attract grants, the quantity and nature of future linkage applications and the potential efficiencies that will arise through the use of more advanced systems.

---

43 Tew, M, Dalziel, KM, Petrie, DJ & Clarke, PM, 2016, ‘Growth of linked hospital data use in Australia, a systematic review’ Australian Health Review, p. 4
However, as WADLS continues to evolve and more processes become automated, these charges are expected to decrease.

The Advisory Group believes that the cost, and the subsequent charges to users, of data linkage in WA must be reduced. In the medium to long term this can be achieved by adopting systems capable of greater automation of administrative checking, streamlined extraction processes (such as CARES) and increasing the number of infrastructure linkages. However, in the short term, greater funding will be required to deliver this.

**FINDING 29**
The funding currently provided to the DLB is not adequate for it to meet the current volume of linkage requests and upgrade its systems to perform linkage more efficiently.

**RECOMMENDATION**
That the State Government increase funding to the Data Linkage Branch. Greater resourcing is essential to support and improve data linkage service provision.
(Recommendation 14)
Chapter 9 – Data Analytics Capability

RECOMMENDATIONS

- State Government agencies invest in data analytics capabilities, and associated training courses for public sector staff be developed. (Recommendation 20)

- The Department of the Premier and Cabinet establish and oversee a unit which coordinates, between Government and the academic sector, cross-agency or whole-of-Government data linkage projects for policy development and implementation. (Recommendation 21)
9.1 The importance of data analytics

Data analytics is about creating meaning from data. It is an essential step in the cycle that involves collecting and managing data, linking data if applicable, analysing data to create insights and knowledge, and applying that knowledge to make a positive difference for the community.

Data analytics can be used for a variety of purposes and in a number of different fields, such as risk analysis, video game design, scientific research, financial modelling, advertising, pricing strategy and business intelligence. Data analytics is an essential tool for the public sector and supports evidence-based decision making in policy development, program management and service delivery.

Whilst data analytics is outside of the scope provided in the ToR for this review, it is noted that the benefits of data linkage are only as good as the analysis performed. Stakeholders providing input into the review also raised related issues, such as: a) the shortfall in the number of skilled analysts and the difficulties in accessing and recruiting appropriately skilled staff; and b) the need for investment and growth in the analytics capability within Government, including increasing the skills of public sector staff and providing access to adequate training. Therefore, data analytics has been considered to some extent in this review, and a number of observations and recommendations have been made.

For the purpose of this report, data analytics using Government data has been described in three categories on the basis that different levels of analytical skills and capabilities are required for different purposes and levels of complexity. Note, in a practical sense there may be different levels within these categories and also some overlap between them. In addition, some agencies may have different needs to what has been outlined below.

1. The most fundamental level is having basic data literacy and an understanding of analytical methods. Arguably the majority of public sector staff, and especially policy and program management staff, should be equipped with these skills.
2. The next level is having the ability to source and manipulate data based on a good understanding of the theoretical and practical aspects of analysis for efficient collection, management, storage and investigation of data. The public sector should maintain professionals with such skills, typically analysts, statisticians or mathematicians, within agencies.
3. At the highest level are scientists and researchers who work with complex, often cross-disciplinary data sets that demand sophisticated skills in data sourcing, modelling, extraction and communication. In general, the public sector should utilise and engage the skills and infrastructure that already exist in this area in the academic and research sectors through partnerships and collaboration.

Every industry, including Government, is poised to benefit from the increasing availability of large volumes of data. Central to harnessing the power of data to drive innovation and informed decision making in the public sector is a workforce that has a broad base of data literacy, embedded data analysts and technical professionals, together with ready access to scientists, researchers and more sophisticated infrastructure and tools.

The level to which these skills and capabilities exist, or are utilised by the State Government, varies between agencies. As WA moves to a whole-of-Government data linkage capacity, this situation needs to be addressed and a workforce that has the skills and capability to analyse and extract the most value out of data is required.

9.2 Data analytics skills and capabilities in the public sector

In a number of submissions to this review, agencies and other stakeholders expressed a need for more data analytic capabilities within the public sector. Stakeholders also noted the shortage of skilled staff within the public sector and the difficulties in recruiting candidates with the required skills. The solutions offered included investment in analytics capabilities within Government, including professional development programs and training courses for public sector staff, and greater engagement with the academic and research community.

The sentiment data capabilities within agencies and the paucity of people with those skills, is consistent with the Australian Government’s Public Sector Data Management Report (December 2015).44

Recommendation 5 of the Public Sector Data Management Report suggested a whole-of-Government approach to building data use and analytics capability within the Australian Government Public Service.

The result was the development of the Australian Public Service Data Skills and Capability Framework as outlined in the publication Data Skills and Capability in the Australian Public Service (August 2016).45

The framework maintains that data skills are essential for all Australian Public Service employees and that there are some roles that have a requirement for more specific (and advanced) data skills, including data analysts, scientists, infrastructure engineers and architects.

Basic data literacy and understanding in the public sector
Data literacy and an understanding of data analysis is a capability that ideally most public service employees should possess, especially those involved in developing and implementing policy, programs and services. At this level, the data user should have an understanding of basic analytical methods, have the ability to extract key messages or underlying trends presented within data, and be able to present results and concepts in a concise and effective manner.

It is expected that there will be a degree of organic growth around these skills in coming years as the use of data across the public sector becomes more prevalent. There is also an increased focus on data capabilities in school-level education through the roll out of the Digital Technologies Curriculum46 and an emergence of a range of external educational activities such as CoderDojo and HackED. In terms of higher education, mathematics, and in some instances statistics, is a requirement for obtaining a science and some business majors.

Nevertheless, additional steps to bolster these capabilities across the WA public sector is strongly advisable. For example, the introduction of foundational and intermediate-level courses in data literacy and applied analytics should be considered. Ideas could be drawn from the initiatives being developed by the Commonwealth Government under its Australian Public Service Data Skills and Capability Framework.

This includes the Data Literacy Programme\textsuperscript{47} which includes learning guides, workshops and e-learning courses to improve general data skills across the Australian Public Service.

\textit{More advanced data skills and capabilities in the public sector}

The ability to source and manipulate data based on a good understanding of the theoretical and practical aspects of analysis for efficient collection, management, storage and investigation of data is a skillset the public sector should enhance within agencies.

Professionals with such skills, typically analysts, statisticians or mathematicians, are able to use different analytical methods and models to identify underlying trends, formulate hypotheses, and test and interpret complex data. Effective data users will also have the ability to extract valid conclusions from various types of data sets that can be used to support decision-making, and can be communicated to a range of technical and non-technical audiences.

There are a number of examples within the WA public sector where agencies are utilising analytical and technical skills to enable data to be collected, shared, linked, analysed and applied to support decision making, policy development and service delivery. This includes, but is not limited to;

- DoH – engages significantly in high level research, but also uses data linkage and analysis to investigate evolving trends, such as in communicable diseases, with the findings informing policy, programs and service delivery;
- Landgate – has developed initiatives such as the SLIP which aggregates Government data in one place for reuse by others. It has 4,000 subscribers across the public, private and research sectors.
- WA Police and the Department of the Attorney General have established a data sharing and linking capacity to support day-to-day operations across these two justice agencies and potentially a third, the Department of Corrective Services.
- Road Safety Commission – is responsible for road safety in WA and analyses crash and injury data to support evidence-based decisions to reduce road trauma.
- Treasury – has a key role in providing the Government with information and advice on economic matters which involves significant data analysis.

The opportunity for the public sector is for all agencies to use data in such a positive fashion. To achieve this, individual agencies, and the State Government more broadly, need to invest in data analytics capabilities, to attract and recruit candidates with these skills and increase the capability of existing staff.

**FINDING 30**
Factors critical to using data to deliver outcomes and services in agencies include:
- executive and managerial support of the use of data as a decision-making tool;
- a foundational level of understanding among staff of the value and limits of data as a decision-making tool; and
- access to analysts, statisticians and data scientists.

It is noted that attracting and recruiting staff with the required data capability may be challenging. As mentioned earlier in this Chapter, the Commonwealth Government’s Public Sector Data Management Report (December 2015)\(^\text{48}\) noted a global undersupply of people with data analytics capabilities. Moreover, some agencies have highlighted difficulties, due to resources, in recruiting the ideal candidates\(^\text{49}\). This could relate to skilled professionals being attracted to other sectors, particularly industry.

It is anticipated this issue will resolve itself to some degree in coming years as the number of individuals who undertake higher education training in mathematics, statistics, data analytics, data science and other related courses increases. This will also be assisted by emerging education and awareness in schools and the community about the importance of data capabilities and the job opportunities in this area.

Furthermore, there has been an increase in postgraduate courses on offer in WA, including UWA’s Master in Data Science and Curtin’s Master of Predictive Analytics (also including a Graduate Certificate and Diploma). However, there is scope for additional courses, such as a Master of Data Analytics, which is currently offered at several other Australian universities.


\(^{49}\) 2016 Data Linkage Review – Submission Process
A necessary complement to recruiting already skilled candidates are courses and professional development opportunities for existing public sector staff. The Commonwealth Government’s publication Data Skills and Capability in the Australian Public Service (August 2016)\(^{50}\) provides a number of examples of such courses offered across Australia, including the following two by UWA:

- Introductory Analysis of Linked Health Data (5 day intensive course)\(^{51}\); and
- Advanced Analysis of Linked Health data (5 day intensive course)\(^{52}\).

It is noted that UWA also offers Introductory Statistics (3 day course)\(^{53}\), to anyone, as well as introducing basic statistics, and R Basics (2 day course)\(^{54}\) for more advanced individuals.

Further consideration would need to be given to the courses available to WA public sector staff, their content suitability, cost and accessibility, and whether additional courses need to be developed.

Finally, if it is not feasible for all individual agencies to maintain in-house data analytics capabilities, consideration could be given to a ‘shared resource’ mechanism whereby staff with the required skills are able to work across agencies to support priority projects, evaluation and policy development.

**RECOMMENDATION**

State Government agencies invest in data analytics capabilities, and associated training courses for public sector staff be developed. (Recommendation 20)

---


\(^{52}\) The University of Western Australia, 2016, ‘School of Population Health – Advanced Analysis of Linked Health Data’, Available from: http://www.sph.uwa.edu.au/courses/winter-spring-summer-school/advanced-analysis [6 November 2016]


9.3 Data analytics for whole-of-Government policy

The most advanced level is working with complex, cross-disciplinary data sets that demand sophisticated skills in data sourcing, modelling, extraction and communication. Data users at this level have the ability to manage large, often interlinked, data sets that require the capacity to oversee secure data storage. They are also able to formulate complex strategies for testing algorithms, identify patterns and mathematical relationships within large volumes of data and understand research involving international standards in analytics.

The skills and infrastructure needed at this level are being developed in the research and academic sectors. These capabilities are also harboured by the private sector, for example the larger finance institutions and mining companies; however, in general this is not easily accessed by the public sector.

WA’s strengths in population data linkage were built on strong partnerships and collaboration between the State Government (DoH), academic sector (UWA and Curtin University) and research institutes (TKI). The long history of this partnership, going back more than 30 years, has resulted in world-leading infrastructure, skills and expertise enabling the use of linked data for research across health and social services.

There is also the potential for greater partnerships and involvement from other universities and research institutes. This relates to using linked data for research across a range of disciplines, but also innovative methods to analyse data.

Some examples include facilitating improved access to data held by other jurisdictions and the Commonwealth Government (CDL at Curtin University); strengthening cyber security measures (ECU Security Research Institute); exploring the use of big data technologies and tools (Pawsey Supercomputing Centre); linking with clinical science and trials (Harry Perkins Institute of Medical Research and Linear Clinical Research); and strengthening use of spatial data and technologies (CRC for Spatial Information).

FINDING 31
A number of WA’s research institutes and universities have a long history of, and significant expertise in, research using linked data, and there is potential for greater involvement from others in future.

The Advisory Group would encourage the public sector to utilise existing capabilities of the research and academic sectors, through partnerships and collaboration.
There are a number of examples of partnerships operating currently between Government and the research sector, in WA and in other jurisdictions, for analysis of linked data that services Government policy and service delivery. This includes but is not limited to:

One-off engagement for a project:

- The TKI was engaged by the DCPFS to undertake a project “Research using linked data to explore outcomes for children who have left out-of-home care”. The partnership enables DCPFS to define the scope of the project and access the expertise within TKI to do the analysis.  

Ongoing engagement for a program of work:

- UWA’s Collaborative for Healthcare Analysis and Statistical Modelling (CHASM) is a project funded by the DoH to provide high level analysis and statistical modelling to inform clinical service planning and service evaluations. The work is directed by the Office of the Chief Medical Officer in the DoH and includes two governance committees comprising a range of stakeholders.

Ongoing engagement and across agencies:

- The DPP was established in 2005 at the TKI and involves fourteen partnering State Government agencies. Research is undertaken by the TKI using linked data from across the partner agencies to better understand the factors affecting the health and social outcomes of WA children and youth. The three-tiered Government structure includes representatives from all partners, and at a range of levels. It has been funded through two consecutive ARC linkage grants and contributions from partners.

There is an opportunity for more of these partnerships to occur, especially with the growth of the State’s data linkage capabilities to a whole-of-Government model and the increase in use of linked data for cross-Government policy solutions.

There are also other models, for example, whereby expertise from the academic and research sector is brought into Government to form a centre, rather than contracting out to the academic and research sector (as per the examples above).

55 2016 Data Linkage Review – Submission Process
56 See note 55
For example, in June 2015, the Data Analytics Centre (DAC) was established within NSW Government, Department of Finance, Services and Innovation to deliver Cabinet-endorsed whole-of-Government data analytics projects. The DAC engages researchers and other external capabilities from universities and other organisations such as CSIRO’s Data61 as needed for certain projects. A cross-disciplinary Advisory Board provides strategic advice to the DAC which is led by the Chief Data Scientist.

To enable the DAC to fulfil its role and remove any barriers to data sharing, the *Data Sharing (Government Sector) Act 2015* became law in November 2015. Also, in South Australia, the *Public Sector (Data Sharing) Bill 2016* was introduced in August 2016 to facilitate the sharing of data between public sector agencies, and public sector agencies and other entities, and to provide for an Office of Data Analytics.

There are a number of benefits associated with this model, including the access legislation, engagement with Cabinet, and the coordination across Government agencies and the research sector on projects.

The Advisory Group notes, however, that data sharing and access legislation could be developed independently of an analytics body, and this may even have a number of benefits, such as data analytics becoming a more widespread tool for Government decision-making. Therefore, the Advisory Group has recommended data sharing legislation be developed regardless (further discussion is provided in Chapter 6 – *Data Availability and Access*).

The Advisory Group also notes the potential resources involved in creating a new physical analytics centre. Whilst creating a hub within Government would have benefits, this may not be entirely necessary if the analysis can be carried out by utilising researchers within academic or research organisations. This has the additional benefit of using existing tools and infrastructure in the research sector. Alternatively, a strong partnership between the Government and research sector such as that of the DPP, but expanded in scope, could be developed.

The Advisory Group notes there are benefits to both models. The mechanism envisaged for data analytics for whole-of-Government policy, is led by Government and engages the analytics tools, skills and infrastructure in the academic and research sectors through close working partnerships. This includes the strategic level to define research priorities and at the working level, in undertaking projects.
FINDING 32
Data analytics expertise is critical for whole-of-Government policy decision-making. There are opportunities to harness the strengths of analytics tools, skills and infrastructure from within the academic and research sectors to enhance Government policy making.

RECOMMENDATION
Government establish and oversee a unit which coordinates, between Government and the academic sector, cross-agency or whole-of-Government data linkage projects for policy development and implementation. (Recommendation 21)
Chapter 10 – Implementation

This report and its recommendations provide a clear path for the future of data linkage, for: (i) short-term improvements to the current data linkage model within the DoH, (ii) medium-term recommendations in the form of privacy legislation, and (iii) longer-term recommendations for the establishment of a Statutory Body to oversee data linkage across Government.

Implementation of these recommendations will be critical, and will require strong leadership through the establishment of a Data Linkage Steering Committee (DLSC) – as indicated through Recommendation 9 in Chapter 3 – Governance. The DLSC will be chaired by the Director General of DPC. Should all recommendations be supported, the DLSC will consist of representatives from DPC, DoH, OGCIO and Treasury as well as an independent member. The DLSC will oversee the following working groups, which will each have a set of recommendations to implement:

1. **Data Linkage Process Improvement and Innovation Working Group** – led by DoH
   - **Membership:** DoH (Chair), DLB, OGCIO, Treasury, Curtin University, UWA, TKI
   - **Role:** Implementation of various recommendations to improve and build upon DLB, DoH data linkage activities.

2. **Whole-of-Government Data Initiatives Working Group** – led by OGCIO
   - **Membership:** OGCIO (Chair), Treasury, DPC, DoH, Landgate, FOI Commission, not-for-profit, industry, research/academia
   - **Role:** Implementation of various recommendations to improve culture and practice of data sharing across Government, engaging with the Not-for-profit and private sectors.

3. **Policy & Legislation Working Group** – led by DPC
   - **Membership:** DPC (Chair), DoH, OGCIO, Treasury, DotAG/SSO, FOI Commission, consumer
   - **Role:** Implementation of recommendations relating to privacy and data sharing legislation.

4. **Data Analytics Working Group** – led by Treasury
   - **Membership:** Treasury (Chair), DPC, OGCIO, DoH, TKI, UWA, Curtin University, Public Sector Commission, other research/academia, other Government agencies
   - **Role:** Implementation of recommendations relating to building whole-of-Government analytics capability.

See Appendix 8 detailing the proposed Data Linkage Steering Committee framework (DLSC).
PART 3 – APPENDICES
Appendix 1 – Terms of Reference

A review of data linkage capabilities in Western Australia to enhance the next generation of whole-of-Government data linkage

This Terms of Reference document sets out the background, scope and intended outcomes of the review of data linkage activities and capabilities in Western Australia (WA). This document will be made public and circulated to stakeholders to assist with consultation and submissions as part of the review. This document also sets out an overview of the review process and the advisory group.

This review has been authorised by Mr Peter Conran, Director General of the Department of the Premier and Cabinet (DPC).

Background and Introduction

It has been well acknowledged that WA has been a world leader in data linkage for many decades, especially in relation to data sets contributing to population health research. Projects and research using linked data have contributed to a range of policy, legislation and investment measures that have improved the health and wellbeing of Western Australians.

However, there is mounting evidence that WA’s position as a leader in data linkage is declining, with reports of linked data users facing barriers, such as long wait times and high costs. These barriers are limiting access to linked data for important health and social welfare research. This in turn reduces the ability to make evidence-based decisions, secure competitive research grants, and attract/retain skilled workers with expertise in data linkage. Ultimately, this results in lost opportunities to improve outcomes and deliver important benefits to citizens.

Data linkage is a technique that connects pieces of information that are thought to relate to the same person, family, place or event. Data linkage can provide many insights, including intervention programs that are working or failing, areas that need greater resources, trends to promote novel intervention strategies, and information regarding risk or preventative factors. Linking data across Government agencies and services can provide a powerful resource for understanding complex issues, as well as developing new multi-disciplinary approaches to improve the WA community.

Over recent years there has been a rapid increase in the demand for linked data for research, Government planning and commercial purposes, with an expansion beyond the health sector. Much of this is in response to a realisation that many “wicked problems” facing modern democracies require a whole-of-Government response (for example: social health determinants, climate change, inequalities, youth incarceration, mental health problems, youth disengagement and educational problems). A whole-of-Government data capacity can enable the best information on these problems and an ability to evaluate whether current or proposed services and policies are working or likely to.

In April 2015 the Premier released A Science Statement for Western Australia (the Science Statement) which recognises data and data intensive science as areas of opportunity for the State to help broaden its economy and create a new generation of jobs. The Science Statement acknowledged WA’s competitive advantages in health and population-level data linkage, together with longitudinal studies such as the Busselton Health Study and Raine Study. One significant commitment under the Science Statement was the development of a new data linkage model that builds on these strengths.
The WA Whole-Of-Government Open Data Policy was also released in 2015. In addition, the WA Government Chief Information Officer (GCIO) was appointed and a supporting office, the Office of the GCIO (OGCIO), was established. These important initiatives form part of the Government’s plan to transform the way public sector data is managed, accessed and used in order to remain competitive in an increasingly digital economy. The OGCIO is working to build the foundations and capabilities necessary to generate a sophisticated and efficient approach to data utilisation across the public sector.

It is, therefore, important and timely to review WA’s current data linkage activities to ensure the State maximises this unique resource. It is essential that this capability meets the growing demand and increasing scope of data linkage, if it is to continue delivering social and economic benefits to Western Australians.

**Scope and intended outcomes**

It is intended that the review of WA’s data linkage capabilities will provide Government with the information necessary to meet the growing demand for data linkage and enhance the next generation of whole-of-Government data linkage for WA.

The key focus areas of the review are to examine:

i. the State’s strengths, capabilities and capacities in data linkage and consider how they can best be utilised and developed;

ii. the barriers and impediments to data linkage, and address how they can be improved;

iii. examine the data linkage systems, infrastructure and processes to determine whether they drive efficient, high quality and secure outcomes; and

iv. examine the governance and funding arrangements for data linkage and assess what is required to meet increasing demand beyond just health.

The review will have regard to:

- current and future requirements for data linkage in WA, including needs, technological advances and trends in service delivery; and
- State and national Government policies, and data linkage models and practices nationally and internationally.

The key outputs of the review are to provide:

- recommendations to Government for immediate improvements to data linkage that will bolster the State’s competitive advantage; and
- options for enhanced data linkage models for the next generation of whole-of-Government data linkage.

**Process and Timing**

The review will be undertaken by the Data Linkage Expert Advisory Group, comprising the WA Chief Scientist, Professor Peter Klinken (Chair), the WA GCIO, Mr Giles Nunis, and Professor Fiona Stanley, Patron and Founding Director of the Telethon Kinds Institute and Professor of Child Health at The University of Western Australia (UWA). More information about members of the advisory group is provided below.
The advisory group will be responsible for developing a report of findings and recommendations against the *scope and intended outcomes* outlined above. The process will include research and consultation across Government and the research sector. The review process will also provide an opportunity for written submissions to be made against the terms of reference of the review.

The advisory group will report to the Director General of the Department of the Premier and Cabinet, with a final report to be submitted to the Director General by August 2016. The report will inform Government on how to proceed, in both the short and longer terms, to meet the growing demand for data linkage and achieve the next generation of whole of Government data linkage for WA.

Given the WA Data Linkage System is managed by the Department of Health, and WA’s strong data linkage foundations in the health sector, Dr David Russell-Weisz, Director General of the Department of Health has endorsed the review and terms of reference.

Information about the review will also be made available on the Department of the Premier and Cabinet website, including opportunities to provide information and submissions: [https://www.dpc.wa.gov.au/Consultation](https://www.dpc.wa.gov.au/Consultation)

**Data Linkage Expert Advisory Group**

The advisory group comprises:

**Professor Peter Klinken (Chair)** – Professor Klinken was appointed to the role of WA Chief Scientist in 2014. Professor Klinken is a leading WA medical research scientist, highly regarded for advancing the understanding of genes involved in leukaemia, cancer and anaemia. His many research achievements include the discovery of a gene that suppresses the growth of tumours.

Professor Klinken’s previous roles have included Professor in Clinical Biochemistry at UWA; Director of Research at the Royal Perth Hospital; and Director of the Harry Perkins Institute of Medical Research (previously the WA Institute for Medical Research). Under his stewardship, the Perkins Institute attracted world-class national and international researchers to WA and made numerous acclaimed medical discoveries. He also led the development of two new state-of-the-art medical research facilities, Perkins North in Nedlands (QEII Medical Centre) and Perkins South in Murdoch (Fiona Stanley Hospital).

**Mr Giles Nunis** – Mr Nunis was appointed the inaugural Chief Executive and Government Chief Information Officer (GCIO) on 30 September 2015. Prior to that, Mr Nunis was the Deputy Director General at the WA Department of State Development. He was responsible for negotiating commercial State Agreements with major resource companies and managing the Government’s international trade offices. During this time Mr Nunis was also seconded to the Department of Health and played a major role in restructuring of the Department’s Fiona Stanley Hospital ICT Program which was successfully delivered on time for the opening of the hospital. Prior to joining the WA Government, Mr Nunis was the National Managing Director of a multi-national ICT and Management Consulting firm, based in Melbourne, with its head office in Zurich, Switzerland.

**Professor Fiona Stanley AC** – Professor Stanley is a Professor of Child Health at UWA, and Patron and Founding Director of the Telethon Kids Institute. Professor Stanley is medically trained with a strong interest in maternal and child health epidemiology and public health.
Professor Stanley has spent much of her career investigating the causes of major childhood illnesses, such as birth defects. One of her major contributions was to establish the Telethon Institute for Child Health Research (now known as the Telethon Kids Institute), an independent multidisciplinary research institute focusing on the causes and prevention of major problems affecting children and youth. Professor Stanley is a passionate advocate for the needs of children and their families, promoting the importance of population data to provide significant health, social and economic benefits to the community. For her research on behalf of Australia’s children and Aboriginal social justice, she was named Australian of the Year in 2003 and was made a UNICEF Australia Ambassador for Early Childhood Development in 2006.

The Office of Science in the Department of the Premier and Cabinet and the Office of the Government Chief Information Officer will support the advisory group in undertaking the review.

Questions and Contact

Information about the review will also be made available on the Department of the Premier and Cabinet website, including opportunities to provide information and submissions: www.dpc.wa.gov.au/Consultation

For additional queries, contact:
Office of Science, Department of the Premier and Cabinet
Email: science@dpc.wa.gov.au
Phone: +61 8 6552 6080
Appendix 2 – Data Linkage Expert Advisory Group

Chair

Professor Peter Klinken, ATSO
Professor Klinken was appointed to the role of WA Chief Scientist in 2014. Professor Klinken is a leading WA medical research scientist, highly regarded for advancing the understanding of genes involved in leukaemia, cancer and anaemia. His many research achievements include the discovery of a gene that suppresses the growth of tumours.

Professor Klinken’s previous roles have included Professor in Clinical Biochemistry at UWA; Director of Research at the Royal Perth Hospital; and Director of the Harry Perkins Institute of Medical Research (previously the WA Institute for Medical Research). Under his stewardship, the Perkins Institute attracted world-class national and international researchers to WA and made numerous acclaimed medical discoveries. He also led the development of two new state-of-the-art medical research facilities, Perkins North in Nedlands (QEII Medical Centre) and Perkins South in Murdoch (Fiona Stanley Hospital).

Members

Professor Fiona Stanley AC
Professor Stanley is a Professor of Child Health at UWA, and Patron and Founding Director of TKI. Professor Stanley is medically trained with a strong interest in maternal and child health epidemiology and public health. Professor Stanley has spent much of her career investigating the causes of major childhood illnesses, such as birth defects. One of her major contributions was to establish the Telethon Institute for Child Health Research (now known as TKI), an independent multidisciplinary research institute focussing on the causes and prevention of major problems affecting children and youth.

Professor Stanley is a passionate advocate for the needs of children and their families, promoting the importance of population data to provide significant health, social and economic benefits to the community. For her research on behalf of Australia’s children and Aboriginal social justice, she was named Australian of the Year in 2003 and was made a UNICEF Australia Ambassador for Early Childhood Development in 2006.
Mr Giles Nunis

Mr Nunis was appointed the inaugural Chief Executive and the WA GCIO on 30 September 2015. Prior to that, Mr Nunis was the Deputy Director General at the WA Department of State Development. He was responsible for negotiating commercial State Agreements with major resource companies and managing the Government’s international trade offices. During this time Mr Nunis was also seconded to the DoH and played a major role in restructuring of the Department’s Fiona Stanley Hospital ICT Program which was successfully delivered on time for the opening of the hospital.

Prior to joining the WA Government, Mr Nunis was the National Managing Director of a multi-national ICT and Management Consulting firm, based in Melbourne, with its head office in Zurich, Switzerland.

Assisting Members

Professor Bruce Armstrong

Professor Armstrong is an Emeritus Professor of Public Health in the Sydney School of Public Health, University of Sydney and Senior Adviser in the Sax Institute Sydney. He has been the Director of the Australian Institute of Health and Welfare, Deputy Director of the International Agency for Research on Cancer, and the Commissioner of Health in WA.

Professor Armstrong is an internationally pre-eminent cancer epidemiologist, acknowledged as a passionate, inspiring leader in cancer research and management. He is an international expert on the causes of skin cancer and melanoma and has made important contributions to knowledge on the causes and control of other cancers, high blood pressure and heart disease. During his career, Professor Armstrong was also the Director of Health Research and Planning in the Health Department of WA, a role which involved working with the State’s earliest linked data.

Professor Louisa Jorm

Professor Louisa Jorm is the Foundation Director of the Centre for Big Data Research in Health at the University of New South Wales. She is an Australian leader in epidemiological research using large-scale and linked health data, including hospital inpatient, cohort study and Medicare data. Her research spans topics including health system performance, evaluation of health services and programs, maternal and child health, and Aboriginal health. She is a high-profile advocate for more and better use of routinely collected health data for research.
Professor Jorm has demonstrated a career commitment to putting evidence to work in policy and practice and her work has had numerous translational impacts, e.g.: driving changes to reporting of national health performance indicators; supporting the planning of community-based early childhood and aged care services; informing national guidelines for management of acute coronary syndrome in Indigenous people; and shaping national policy regarding access to publicly funded health data for research.

She has played a leading role in the establishment of major infrastructure and capacity for health big data research in Australia, including the NSW/ACT CHeReL, the 45 and Up Study and the NSW Biostatistical Officer Training Program. She led the development of SURE, a facility that benefits researchers nationally by providing secure remote access to linked health data. In the last five years Professor Jorm has published more than 60 scientific papers, been awarded more than $11M in research funding and given invited plenary presentations at five international and eleven national conferences.
# Appendix 3 – Information Session Attendees

## Data Linkage Review Information Session Attendees

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mr Satch Kirkpatrick</td>
<td>Director</td>
<td>Adaptive Process Design Pty Ltd</td>
</tr>
<tr>
<td>2 Mrs Zaneta Georgievski</td>
<td>Strategic Partnerships Officer</td>
<td>Australian Bureau of Statistics</td>
</tr>
<tr>
<td>3 Miss Renae Smith</td>
<td>SPM</td>
<td>Australian Bureau of Statistics</td>
</tr>
<tr>
<td>4 Mr Darren Mottolini</td>
<td>WA Business and Research Manager</td>
<td>CRC for Spatial Information</td>
</tr>
<tr>
<td>5 Prof. James Boyd</td>
<td>Director, Centre for Data Linkage</td>
<td>Curtin University</td>
</tr>
<tr>
<td>6 Prof. Anna Ferrante</td>
<td>Deputy Director, PHRN Centre for Data Linkage</td>
<td>Curtin University</td>
</tr>
<tr>
<td>7 Mr Paul Nicholls</td>
<td>Director Strategic Projects (R&amp;D)</td>
<td>Curtin University</td>
</tr>
<tr>
<td>8 Prof. James Semmens</td>
<td>Chair, Health Innovation</td>
<td>Curtin University</td>
</tr>
<tr>
<td>9 Ms Sarah French</td>
<td>Manager Research and Evaluation</td>
<td>Department for Child Protection and Family Support</td>
</tr>
<tr>
<td>10 Ms Samantha Kelly</td>
<td>Senior Performance Analyst</td>
<td>Department of Corrective Services</td>
</tr>
<tr>
<td>11 Mr Dale Miller</td>
<td>Manager, System Performance</td>
<td>Department of Education</td>
</tr>
<tr>
<td>12 Mr Mark Almasi</td>
<td>Assistant Director, Intelligence &amp; Investigations</td>
<td>Department of Finance</td>
</tr>
<tr>
<td>13 Mrs Kylie Corbett</td>
<td>Manager Information and Systems</td>
<td>Department of Finance</td>
</tr>
<tr>
<td>14 Mr Andy Wood</td>
<td>Executive Director Corporate Services</td>
<td>Department of Finance</td>
</tr>
<tr>
<td>15 Mr Brendan Power</td>
<td>Chief Information Officer</td>
<td>Department of Fisheries</td>
</tr>
<tr>
<td>Name</td>
<td>Position</td>
<td>Organisation</td>
</tr>
<tr>
<td>---------------------------</td>
<td>----------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>Dr Janine Alan</td>
<td>Program Manager, Data Linkage Branch</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Mr Peter Cosgrove</td>
<td>CARES Data Co-ordinator</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Ms Alexandra Godfrey</td>
<td>Project Manager, DLB</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Prof Gary Geelhoed</td>
<td>Chief Medical Officer</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Dr Laura Miller</td>
<td>Epidemiologist</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Mr Tim Reid</td>
<td>Group Director Performance</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Mr Tony Satti</td>
<td>Director Data Integrity</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Dr Debbie Turner</td>
<td>Manager, Office of the Chief Medical Officer</td>
<td>Department of Health</td>
</tr>
<tr>
<td>Ms Jennifer Endersbee</td>
<td>Principal Statistical Analyst</td>
<td>Department of the Attorney General</td>
</tr>
<tr>
<td>Ms Melissa Rudez</td>
<td>Principal Policy Officer</td>
<td>Department of the Premier and Cabinet</td>
</tr>
<tr>
<td>Ms Sally Whatmough</td>
<td>Principal Policy Officer</td>
<td>Department of the Premier and Cabinet</td>
</tr>
<tr>
<td>Mrs Hareg Ambaye</td>
<td>Manager Research and GIS</td>
<td>Department of Training and Workforce Development</td>
</tr>
<tr>
<td>Mr Mark Bloomfield</td>
<td>Director, Performance Evaluation and Statistics</td>
<td>Department of Training and Workforce Development</td>
</tr>
<tr>
<td>Mr Michael Eckermann</td>
<td>Manager, Economic and Labour Market Analysis</td>
<td>Department of Training and Workforce Development</td>
</tr>
<tr>
<td>Mr Greg Hilton</td>
<td>Senior Labour Market Analyst</td>
<td>Department of Training and Workforce Development</td>
</tr>
<tr>
<td>Mr Geoff Robertson</td>
<td>Manager</td>
<td>Department of Training and Workforce Development</td>
</tr>
<tr>
<td>Name</td>
<td>Position</td>
<td>Organisation</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>32 Mr Simon Walker</td>
<td>Executive Director</td>
<td>Department of Training and Workforce Development</td>
</tr>
<tr>
<td>33 Ms Natalia Kacperek</td>
<td>Business Intelligence Analyst</td>
<td>Department of Transport</td>
</tr>
<tr>
<td>34 Mr Luke Crotty</td>
<td>Senior Analyst</td>
<td>Department of Treasury</td>
</tr>
<tr>
<td>35 Mrs Kylie Dalling</td>
<td>Manager</td>
<td>Department of Treasury</td>
</tr>
<tr>
<td>36 Mr Coan Harvey</td>
<td>Manager</td>
<td>Department of Treasury</td>
</tr>
<tr>
<td>37 Mr Kurt Simba</td>
<td>Director</td>
<td>Department of Treasury</td>
</tr>
<tr>
<td>38 Ms Annie MacKinnon</td>
<td>Consumer Representative DPP</td>
<td>Developmental Pathways Project Community Reference Group</td>
</tr>
<tr>
<td>39 Mr Vaughan Bevan</td>
<td>Planning and Research Officer</td>
<td>Disability Services Commission</td>
</tr>
<tr>
<td>40 Kate Clarke</td>
<td></td>
<td>Disability Services Commission</td>
</tr>
<tr>
<td>41 Mr Harnidh Singh</td>
<td>Principal Planning &amp; Research Analyst</td>
<td>Disability Services Commission</td>
</tr>
<tr>
<td>42 Dr Agi Gedeon</td>
<td>Research Data Consultant, Australian National Data Service (ANDS)</td>
<td>Edith Cowan University</td>
</tr>
<tr>
<td>43 Dr Darren Gibson</td>
<td>Collaboration and Innovation Manager</td>
<td>Edith Cowan University</td>
</tr>
<tr>
<td>44 Mrs Jillian Scott</td>
<td>Manager ICT</td>
<td>Forest Products Commission</td>
</tr>
<tr>
<td>45 Mr Mark Caldwell</td>
<td>GM Finance &amp; Business Services</td>
<td>GESB</td>
</tr>
<tr>
<td>46 Ms Pip Brennan</td>
<td>Executive Director</td>
<td>Health Consumers' Council</td>
</tr>
<tr>
<td>47 Mr Carmelo Naso</td>
<td>Chief information officer</td>
<td>Mainroads WA</td>
</tr>
<tr>
<td>48 Mr Sven Bluemmel</td>
<td>Information Commissioner</td>
<td>Office of the Information Commissioner</td>
</tr>
<tr>
<td>Name</td>
<td>Position</td>
<td>Organisation</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Ms Su Lloyd</td>
<td>Principal Legal Officer</td>
<td>Office of the Information Commissioner</td>
</tr>
<tr>
<td>Mrs Elyse Loughton</td>
<td>Paralegal Officer</td>
<td>Office of the Information Commissioner</td>
</tr>
<tr>
<td>Mr David Zarb</td>
<td>CEO</td>
<td>PlaygroupWA</td>
</tr>
<tr>
<td>Dr Felicity Flack</td>
<td>Manager Policy and Client Services</td>
<td>Population Health Research Network</td>
</tr>
<tr>
<td>Ms Ingrid Landwehr</td>
<td>Manager Contracts and Finance</td>
<td>Population Health Research Network</td>
</tr>
<tr>
<td>Mr Jeff Spence</td>
<td>Senior Business Analyst</td>
<td>Road Safety Commission</td>
</tr>
<tr>
<td>Mrs Jenny Offer</td>
<td>Principal Consultant-Certification</td>
<td>School Curriculum and Standards Authority</td>
</tr>
<tr>
<td>Mr Sean Chinnery</td>
<td>General Manager, Performance &amp; Planning</td>
<td>St John Ambulance (WA)</td>
</tr>
<tr>
<td>Mr Ashley Morris</td>
<td>Technical Services Director</td>
<td>St John Ambulance (WA)</td>
</tr>
<tr>
<td>Dr Lyn Colvin</td>
<td>Program Manager, The ORIGINS Project</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>Prof Nick de Klerk</td>
<td>Head of Biostatistics</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>Dr Rebecca Glauert</td>
<td>Program Manager, Developmental Pathways Project</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>Dr Michele Hansen</td>
<td>NHMRC Early Career Fellow</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>Dr Helen Leonard</td>
<td>Principal Research Fellow</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>Dr Carrington Shepherd</td>
<td>Research Fellow</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>Dr Bahareh Badrian</td>
<td>Project Manager, Collaborative Research Network</td>
<td>The University of Notre Dame</td>
</tr>
<tr>
<td>Prof Matthew Knuiman</td>
<td>Professor, Population Health</td>
<td>The University of Western Australia</td>
</tr>
<tr>
<td>Name</td>
<td>Position</td>
<td>Organisation</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>66 Dr David Lawrence</td>
<td>Principal Research Fellow</td>
<td>The University of Western Australia</td>
</tr>
<tr>
<td>67 Prof David Preen</td>
<td>Chair in Public Health</td>
<td>The University of Western Australia</td>
</tr>
<tr>
<td>68 Dr David Whyatt</td>
<td>Senior Research Fellow</td>
<td>The University of Western Australia</td>
</tr>
<tr>
<td>69 Prof John Challis</td>
<td>Executive Director</td>
<td>WA Health Translation Network</td>
</tr>
<tr>
<td>70 Mr Les Bechelli</td>
<td>Director Business Strategy &amp; Finance</td>
<td>WA Police</td>
</tr>
<tr>
<td>71 Mrs Aleisha Godenzie</td>
<td>Change Manger - Business Intelligence Office</td>
<td>WA Police</td>
</tr>
<tr>
<td>72 Mr Bruce Miller</td>
<td>Assistant Director Business Intelligence</td>
<td>WA Police</td>
</tr>
<tr>
<td>73 Mr Chris Twomey</td>
<td>Director, Social Policy</td>
<td>WACOSS</td>
</tr>
<tr>
<td>74 Mr Danny Murphy</td>
<td>Manager - Digital Services</td>
<td>Western Australian Museum</td>
</tr>
</tbody>
</table>

**Date Linkage Expert Advisory Group Attendees**

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Professor Peter Klinken</td>
<td>Chair</td>
</tr>
<tr>
<td>2 Professor Fiona Stanley</td>
<td>Member</td>
</tr>
<tr>
<td>3 Professor Bruce Armstrong</td>
<td>Assisting Member</td>
</tr>
</tbody>
</table>

**Speakers**

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mr Peter Conran</td>
<td>Director General</td>
<td>Department of the Premier and Cabinet</td>
</tr>
<tr>
<td>2 Ms Fiona Roche</td>
<td>Executive Director</td>
<td>Office of Science</td>
</tr>
<tr>
<td>3 Mr Damian Shepherd</td>
<td>Director</td>
<td>Landgate</td>
</tr>
<tr>
<td>4 Professor Desiree Silva</td>
<td>Head of Paediatrics</td>
<td>Joondalup Health Campus</td>
</tr>
</tbody>
</table>
Support Staff – Office of Science, Department of the Premier and Cabinet

<table>
<thead>
<tr>
<th></th>
<th>Name</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ms Kathryn James</td>
<td>Director</td>
</tr>
<tr>
<td>2</td>
<td>Ms Karina Schaap</td>
<td>Senior Policy Officer</td>
</tr>
<tr>
<td>3</td>
<td>Ms Amanda Low</td>
<td>Graduate Officer</td>
</tr>
<tr>
<td>4</td>
<td>Mr Adam Byass</td>
<td>Graduate Officer</td>
</tr>
<tr>
<td>5</td>
<td>Mr Jacob Goldschlager</td>
<td>Senior Project Officer</td>
</tr>
</tbody>
</table>
## Appendix 4 – Submissions Received

<table>
<thead>
<tr>
<th></th>
<th>Individual/Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bridgette McNamara on behalf of the research team for the Defying the Odds project</td>
</tr>
<tr>
<td>2</td>
<td>Cooperative Research Centre for Spatial Information</td>
</tr>
<tr>
<td>3</td>
<td>Curtin University</td>
</tr>
<tr>
<td>4</td>
<td>Curtin University – PHRN Centre for Data Linkage</td>
</tr>
<tr>
<td>5</td>
<td>David Lydon (individual submission)</td>
</tr>
<tr>
<td>6</td>
<td>Department for Child Protection and Family Support</td>
</tr>
<tr>
<td>7</td>
<td>Department of Aboriginal Affairs</td>
</tr>
<tr>
<td>8</td>
<td>Department of Education</td>
</tr>
<tr>
<td>9</td>
<td>Department of Health</td>
</tr>
<tr>
<td>10</td>
<td>Department of Health – Communicable Disease Control Directorate</td>
</tr>
<tr>
<td>11</td>
<td>Department of the Attorney General</td>
</tr>
<tr>
<td>12</td>
<td>Department of Training and Workforce Development</td>
</tr>
<tr>
<td>13</td>
<td>Department of Treasury</td>
</tr>
<tr>
<td>14</td>
<td>Edith Cowan University</td>
</tr>
<tr>
<td>15</td>
<td>Hannah Moore on behalf of the Infectious Diseases Epidemiology Team, Wesfarmers Centre for Vaccines and Infectious Diseases, Telethon Kids Institute.</td>
</tr>
<tr>
<td>16</td>
<td>Harry Perkins Institute</td>
</tr>
<tr>
<td>17</td>
<td>Janice Wong on behalf of students at the Telethon Kids Institute</td>
</tr>
<tr>
<td>18</td>
<td>Judy Allen (individual submission)</td>
</tr>
<tr>
<td>19</td>
<td>Landgate and Department of Mines and Petroleum</td>
</tr>
<tr>
<td>20</td>
<td>Linear Clinical Research and Collaborators</td>
</tr>
<tr>
<td></td>
<td>Individual/Organisation</td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>21</td>
<td>Lions Eye Institute</td>
</tr>
<tr>
<td>22</td>
<td>Main Roads WA</td>
</tr>
<tr>
<td>23</td>
<td>Mental Health Commission</td>
</tr>
<tr>
<td>24</td>
<td>Merran Smith (individual submission)</td>
</tr>
<tr>
<td>25</td>
<td>NSW Ministry of Health – Centre for Health Record Linkage</td>
</tr>
<tr>
<td>26</td>
<td>Office of the Information Commissioner</td>
</tr>
<tr>
<td>27</td>
<td>Playgroup WA</td>
</tr>
<tr>
<td>28</td>
<td>Population Health Research Network</td>
</tr>
<tr>
<td>29</td>
<td>Raine Study</td>
</tr>
<tr>
<td>30</td>
<td>Road Safety Commission</td>
</tr>
<tr>
<td>31</td>
<td>St John of God Hospital Subiaco</td>
</tr>
<tr>
<td>32</td>
<td>State Library of Western Australia</td>
</tr>
<tr>
<td>33</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>34</td>
<td>The University of Notre Dame</td>
</tr>
<tr>
<td>35</td>
<td>The University of Western Australia</td>
</tr>
<tr>
<td>36</td>
<td>The University of Western Australia – Collaborative Health Analysis and Statistical Modelling</td>
</tr>
<tr>
<td>37</td>
<td>The University of Western Australia – Faculty of Education</td>
</tr>
<tr>
<td>38</td>
<td>The University of Western Australia – School of Population Health</td>
</tr>
<tr>
<td>39</td>
<td>The University of Western Australia and Telethon Kids Institute – Consumer &amp; Community Involvement Program</td>
</tr>
<tr>
<td>40</td>
<td>WA Council of Social Service</td>
</tr>
<tr>
<td>41</td>
<td>WA Police</td>
</tr>
</tbody>
</table>
## Appendix 5 – Consultation Meetings

The following meetings were held with members of Advisory Group as well as support staff:

<table>
<thead>
<tr>
<th>Date</th>
<th>Name</th>
<th>Position / Area</th>
<th>Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Aug 2016</td>
<td>Merran Smith</td>
<td>Chief Executive</td>
<td>Population Health Research Network</td>
</tr>
<tr>
<td></td>
<td>Felicity Flack</td>
<td>Manager, Policy and Client Services</td>
<td></td>
</tr>
<tr>
<td>3 Aug 2016</td>
<td>Anne McKenzie</td>
<td>Manager, Consumer &amp; Community Involvement Program</td>
<td>The University of Western Australia and Telethon Kids Institute</td>
</tr>
<tr>
<td></td>
<td>David Lawrence</td>
<td>Principal Research Fellow, Graduate School of Education</td>
<td>The University of Western Australia</td>
</tr>
<tr>
<td></td>
<td>David Preen</td>
<td>Chair in Public Health</td>
<td></td>
</tr>
<tr>
<td></td>
<td>David Whyatt</td>
<td>Senior Research Fellow</td>
<td></td>
</tr>
<tr>
<td>8 Aug 2016</td>
<td>Judith (Judy) Allen</td>
<td>Judy provided an individual submission into the review. Her experience as a legal academic at The University of Western Australia and Chair of the Department of Health Human Research Ethics Committee was noted.</td>
<td></td>
</tr>
<tr>
<td>8 Aug 2016</td>
<td>Kurt Sibma</td>
<td>A/Director, Performance &amp; Evaluation Group 3</td>
<td>Department of Treasury</td>
</tr>
<tr>
<td></td>
<td>Kylie Dalling</td>
<td>Manager, Program Evaluation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Luke Crotty</td>
<td>Senior Analyst</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Helen Ensikat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Aug 2016</td>
<td>Damien Shepherd</td>
<td>Director, WA Land Information System</td>
<td>Landgate</td>
</tr>
<tr>
<td></td>
<td>Neville D’Antoine</td>
<td>A/General Manager, Geoscience Information Geological Survey</td>
<td>Department of Mines and Petroleum</td>
</tr>
<tr>
<td>Date</td>
<td>Name</td>
<td>Position / Area</td>
<td>Organisation</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------</td>
<td>----------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>8 Aug 2016</td>
<td>Jonathon Carapetis</td>
<td>Director</td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td></td>
<td>Nick de Klerk</td>
<td>Head of Biostatistics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rebecca Glauert</td>
<td>Head of Data Linkage and Social Policy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Judy Katzenellenbogen</td>
<td>Head of Rheumatic Heart Disease Public Health Research, Infections and Vaccines</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hannah Moore</td>
<td>Head of Infectious Disease and Epidemiology</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Michelle Hansen</td>
<td>NHMRC Early Career Fellow</td>
<td></td>
</tr>
<tr>
<td>15 Aug 2016</td>
<td>Andrew Marshall</td>
<td>Manager, Legal Policy and Analysis</td>
<td>Department of the Attorney General</td>
</tr>
<tr>
<td></td>
<td>Jennifer Endersbee</td>
<td>Principal Statistical Officer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monica Cass</td>
<td>Senior Evaluation and Research Officer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bruce Miller</td>
<td>Assistant Director, Business Intelligence and Analytics</td>
<td>WA Police</td>
</tr>
<tr>
<td></td>
<td>David Lydon</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aleisha Godenzie</td>
<td>Change Manger – Business Intelligence Office</td>
<td></td>
</tr>
<tr>
<td>15 Aug 2016</td>
<td>James Boyd</td>
<td>Director, Centre for Data Linkage</td>
<td>Curtin University</td>
</tr>
<tr>
<td>Date</td>
<td>Name</td>
<td>Position / Area</td>
<td>Organisation</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------</td>
<td>------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>15 Aug 2016</td>
<td>Anna Ferrante</td>
<td>Deputy Director, Centre for Data Linkage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chris Twomey</td>
<td>Director, Social Policy</td>
<td>WA Council of Social Service</td>
</tr>
<tr>
<td></td>
<td>Penny Knight</td>
<td></td>
<td>Curtin University</td>
</tr>
<tr>
<td></td>
<td>David Ansell</td>
<td></td>
<td>Telethon Kids Institute</td>
</tr>
<tr>
<td>15 Aug 2016</td>
<td>Tarun Weeramanthri</td>
<td>Assistant Director General, Public Health</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tony Satti</td>
<td>Acting Director, Data Collections, Information Data &amp; Standards</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Janine Alan</td>
<td>Program Manager, Data Linkage Branch, Information Data &amp; Standards</td>
<td>Department of Health</td>
</tr>
<tr>
<td></td>
<td>Alex Godfrey</td>
<td>Project Manager, Information Data &amp; Standards</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Robyn Daniels</td>
<td>Senior Legal Advisor</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Steve Ashburn</td>
<td>Senior Solicitor</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 6 – Data Linkage Framework

**Data Collection**
- Collected by organisations

**Data Linkage**
- Performed by DoH and other organisations
  - See Appendix 7 for DoH process, which includes the following steps:
    1. Application
    2. Approvals
    3. Part 1 – Linkage
    4. Extraction
    5. Part 2 – Linkage and QA

**Data Analytics**
- Analysing data to make meaning
  - Linked data is analysed in WA to support approved research and has the potential to support development of policies, and also for the planning and evaluation of services.
  - Three bands of Data Analytics:
    1. Capacity to understand data analysis, its limitations, uses and ability to apply to achieve policy outcomes.
    2. Ability to use data to evaluate its effectiveness.
    3. Ability to research complex, often cross-disciplinary issues that demand sophisticated and highly competent data analysts.

**Benefits**
- Delivering and evaluating
  - Ability to make more informed policy decisions and provide services more effectively and efficiently.
  - Examples include:
    - Improve strategic policy, programs and services
    - Improve welfare outcomes, e.g. housing, disability
    - Develop early intervention and diversion programs for the criminal justice system
    - Deliver better road and transport systems
    - Aid service delivery in improving health outcomes in remote and regional areas
    - Improve educational outcomes
Appendix 7 – Data Linkage Process

Overview of data linkage process

1. APPLICATION
   - DLB, Researcher and Data Custodian/s work together to develop the draft application
   - Data Custodian/s provide in-principle support

2. APPROVALS
   - DoH HREC approval
   - Data Custodian/s approval

3. EXTRACTION
   - DLB gathers raw data from Data Custodian/s, creates linkage keys then sends to Data Custodian/s

4. LINKAGE & QA
   - Data Custodian/s attach content data to linkage keys and sends to DLB
   - DLB merges linkage keys to create the linked data set, quality assures and encrypts
   - DLB provides fully linked data set/s to Researcher

5. Part 2
   - LINKAGE & QA

DE-IDENTIFICATION

DLB ADMINISTRATION OF THE DATA LINKAGE PROCESS
Appendix 9 – Bibliography

2016 Data Linkage Review Submissions – refer to Appendix 4


Coghlan, R, Lawrence, D, Holman, D, Jablensky, A 2001, ‘Duty to Care: Physical illness in people with mental illness’. Perth: The University of Western Australia

‘Chair in Public Health: 20 Years of Discovery and Advance 1994-95 to 2013-14’, 2014. School of Population Health, The University of Western Australia, pp 31


Department of Health, July 2012, HREC Terms of Reference, Available from: [link]
[17 November 2016]


Government of Western Australia, 2016, ‘Media Statements’, ‘Healthy Research Outlook at Science on the Swan, Available from: [link] [17 November, 2016]


Tew, M, Dalziel, KM, Petrie, DJ & Clarke, PM, 2016, ‘Growth of linked hospital data use in Australia, a systematic review’ Australian Health Review, p. 4


