



Australian Government  
The Treasury

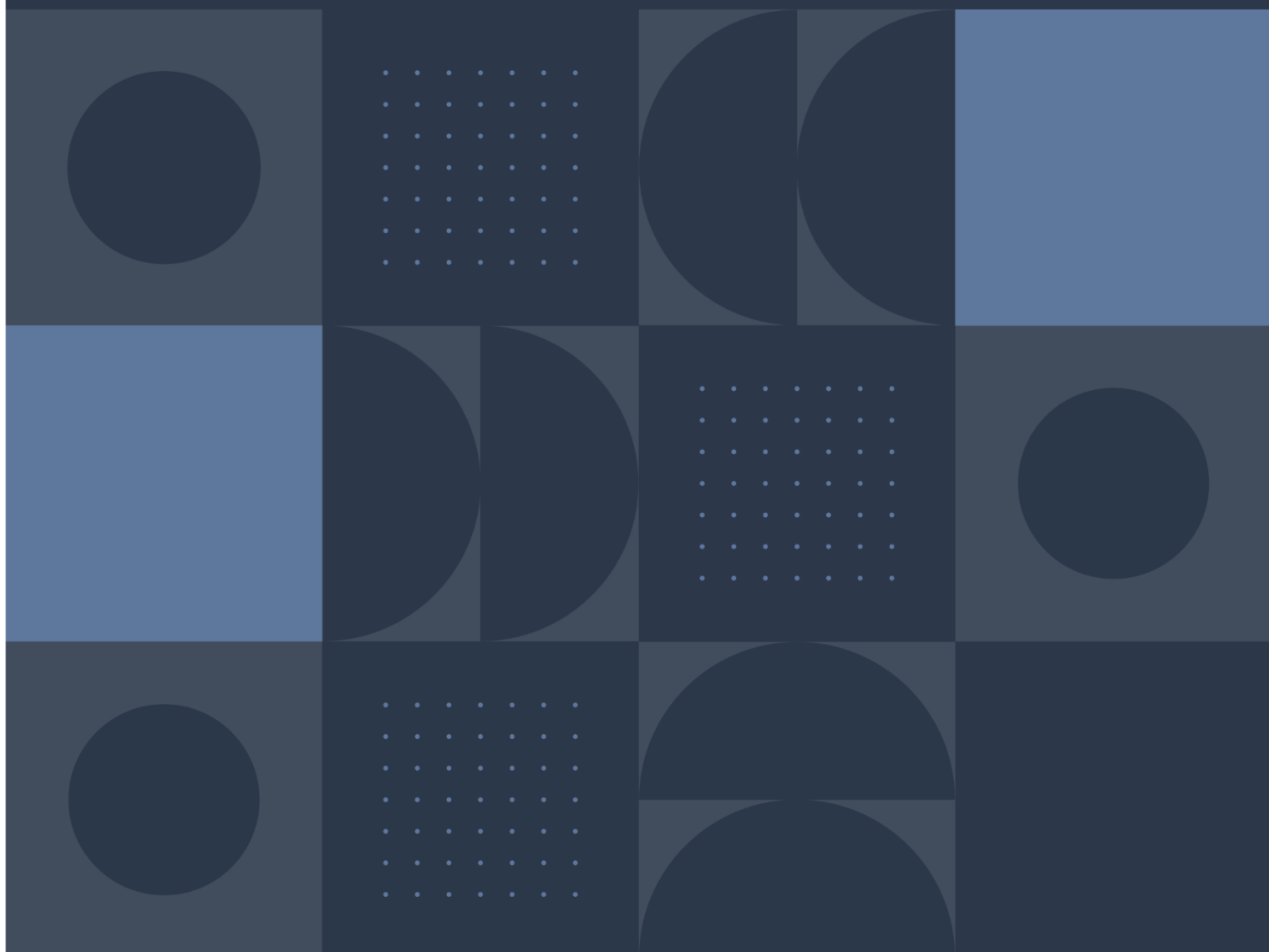
Australian  
Centre for  
Evaluation

**ace**

# Government administrative data sources for evaluation in Australia

A guide prepared by the  
Australian Centre for Evaluation

January 2025



© Commonwealth of Australia 2025

This publication is available for your use under a [Creative Commons Attribution 4.0 International](#) licence, with the exception of the Commonwealth Coat of Arms, the Treasury logo, photographs, images, third party materials, materials protected by a trademark, signatures and where otherwise stated. The full licence terms are available from [creativecommons.org/licenses/by/4.0/legalcode](https://creativecommons.org/licenses/by/4.0/legalcode).



Use of Treasury material under a [Creative Commons Attribution 4.0 International](#) licence requires you to attribute the work (but not in any way that suggests that the Treasury endorses you or your use of the work).

### **Treasury material used ‘as supplied’**

Provided you have not modified or transformed Treasury material in any way including, for example, by changing the Treasury text; calculating percentage changes; graphing or charting data; or deriving new statistics from published Treasury statistics – then Treasury prefers the following attribution:

*Source: The Australian Government the Treasury.*

### **Derivative material**

If you have modified or transformed Treasury material, or derived new material from those of the Treasury in any way, then Treasury prefers the following attribution:

*Based on The Australian Government the Treasury data.*

### **Use of the Coat of Arms**

The terms under which the Coat of Arms can be used are set out on the Department of the Prime Minister and Cabinet website (see [www.pmc.gov.au/government/commonwealth-coat-arms](http://www.pmc.gov.au/government/commonwealth-coat-arms)).

### **Other uses**

Enquiries regarding this licence and any other use of this document are welcome at:

Manager  
Media Unit  
The Treasury  
Langton Crescent  
Parkes ACT 2600  
Email: [media@treasury.gov.au](mailto:media@treasury.gov.au)

# Contents

- Overview and purpose .....1
- Background.....2
- Person Level Integrated Data Asset (PLIDA) .....3
- Business Longitudinal Analysis Data Environment (BLADE) .....6
- ATO Longitudinal Information Files (ALife) .....7
- Longitudinal datasets .....8
- Australian Institute of Family Studies (AIFS) .....11
- The Australian Data Archive (ADA).....12
- Australian Institute of Health and Welfare (AIHW) .....13
- Department of Social Services (DSS) .....15
- State and territory data agencies.....16

## Overview and purpose

This note provides a guide to Government administrative data sources for evaluation in Australia, as at January 2025. It lists the major sources of linked administrative data, major longitudinal surveys, and state and territory data agencies, but this guide is not comprehensive. Buyer beware!

This note was prepared by the Australian Centre for Evaluation (ACE). If you have comments or suggestions on how it could be improved or corrected, please contact us at [evaluation@treasury.gov.au](mailto:evaluation@treasury.gov.au).

# Background

There has been increased interest in recent times to use or repurpose routinely collected administrative datasets for evaluations, in particular for randomised controlled trials (RCT).<sup>1</sup> This is due to the demand for simplified data collection within trials to increase efficiency and reduce costs.

The Australian Bureau of Statistics defines administrative data as information that government departments, businesses and other organisations collect for reasons other than research, such as registrations, billing and record keeping.

There are several advantages for using administrative datasets alongside data collected from RCTs.

First, administrative data may be more reliable as there is potential to eliminate participant recall bias, or other issues arising from self-reporting by participants.

Second, administrative data can serve to provide more accurate and complete data about the participants than survey responses from participants. This is particularly useful when the attrition rates are likely to be higher.

Third, administrative data can be useful in measuring interventions in cluster RCTs as they can allow for extensive coverage of participants without requiring extensive participant-level data collection.

Thus, linked administrative data can provide more comprehensive information about individuals and communities. In addition, linked administrative data is often less expensive than new data collection.

All evaluations need to be robust, ethical, and culturally appropriate. For more information on evaluations of policies and programs affecting Aboriginal and Torres Strait Islander people, see [Indigenous evaluation](#) on Australian Centre for Evaluation (ACE) website.

The Framework for Governance of Indigenous Data provides a stepping stone towards greater awareness and acceptance by Australian Government agencies of the principles of Indigenous Data Sovereignty. The Framework aims to provide Aboriginal and Torres Strait Islander people greater agency over how their data is governed within the Australian Public Service, so government held data better reflects their priorities and aspirations. See the [Framework for Governance of Indigenous Data](#) for more details.

This note details the different administrative data sources available, the datasets within these sources and how to access these data sources.

---

1 Fahridin, S., Agarwal, N., Bracken, K., Law, S., Morton, R. L. (2024), The use of linked administrative data in Australian randomised controlled trials: A scoping review, *Clinical Trials*, doi:10.1177/17407745231225618

# Person Level Integrated Data Asset (PLIDA)

PLIDA, previously known as Multi-Agency Data Integration Project (MADIP), is a person-level secure microdata asset constructed by the Australian Bureau of Statistics (ABS) combining information on health, education, government payments, income and taxation, employment, and population demographics (including the Census) over time. The Productivity Commission has recently published a research paper providing an overview of PLIDA from a user perspective.<sup>2</sup> It takes a case study approach to identifying the strengths and limitations of PLIDA for different areas of Australian public policy research and reporting. It also notes emerging directions that stand to enable greater use by researchers over time.

## Datasets

The PLIDA Modular Product (PMP) is a detailed microdata product created from the PLIDA data asset. All access to the PLIDA data asset is via the PMP. The PMP contains data from a range of different data custodians and covers several domains. See the [Data Item List \(DIL\)](#) for details on all available datasets and the data custodians. See below for a summary of some of the datasets.

### Data Over Multiple Individual Occurrences (DOMINO)

DOMINO contains modular event-based data on individuals providing a longitudinal picture of the interaction of individual welfare recipients throughout their interactions with Department of Social Services (DSS) payments. DOMINO contains information on recipients' demographics, benefits history, concessions, education (where available) and housing. DOMINO data is also available through Australian Institute of Health and Welfare (AIHW). To access this data, you will need to submit all relevant documentation from the [application pack](#) for review by the AIHW.

### Data Exchange (DEX)

The DEX is the program performance reporting solution developed by the DSS. It was developed to improve the way administrative data are collected and used to improve the wellbeing of people and families. Organisations funded under the National Disability Advocacy Program (NDAP) are required under their agreement with the Commonwealth to enter data into the DSS DEX in accordance with the data exchange protocols.

### Census of Population and Housing (Census)

The [Census](#) is conducted every 5 years to measure the number of people in Australia on Census night, their characteristics and the households and dwellings in which they live. By collecting lots of information in a standardised way, from the country as a whole, the Census provides a rich and detailed snapshot of Australia. Australia's 17th (and most recent) national Census was conducted on 10 August 2021.

---

2 Productivity Commission (2024), A-PLIDA-nalysis: Using PLIDA for public policy research and reporting, Research paper, Canberra

## Australian Census Longitudinal Dataset (ACLD)

The [ACLD](#) uses data from the Census of Population and Housing to build a rich research tool for exploring how Australian society changes over time. It can provide new insights into the dynamics and transitions that drive social and economic change over time, and how these vary for diverse population groups and geographies. Four waves of Census data have so far contributed to the ACLD from the 2006 Census (wave 1), 2011 Census (wave 2), 2016 Census (wave 3) and 2021 Census (wave 4).

## Single Touch Payroll (STP)

The Australian Taxation Office (ATO) receives payroll information from employers through STP enabled payroll and accounting software each time the employer runs its payroll. The ATO provides selected employer and job level data items from the STP system to the ABS to produce official statistics. STP data are compiled into weekly indexes of payroll jobs and wages.

## Apprentice and trainee

[Apprentice and trainee](#) collection provides data on all persons employed under a training contract and includes both apprentices and trainees. Data are collected from state training authorities on training activity in apprenticeships and traineeships in Australia, including information on training rates and duration of training.

## Australian Early Development Census (AEDC)

The [AEDC](#) is a national assessment conducted every 3 years to examine how children have developed by the time they start school. It looks at 5 areas of early childhood development: physical health and wellbeing, social competence, emotional maturity, language, and cognitive skills (school-based), and communication skills and general knowledge. Data is collected by teachers of children in their first year of school. Teachers respond to around 100 questions that measure early childhood development across the 5 areas. Children are allocated a score against the domains to determine whether they are developmentally on track, at risk or vulnerable.

## Access

PLIDA is available in [DataLab](#), a secure data analysis environment. Access is limited to government employees/contractors, academics, public policy research institutes and individuals sponsored by government. External organisations can still have access with some limitations.

Users need to meet ABS [safe people criteria](#), including researchers who:

- have completed all relevant undertakings and declarations.
- can use at least one of the statistical analytical languages available in the DataLab.
- have at least 3 years of either quantitative research experience or university study with a significant component working with quantitative data, or have a referral from an experienced researcher working on the same project.
- have an approved safe project that is for statistical and/or research purposes.

- have completed [ABS safe researcher training](#) (current wait time is approximately 4 weeks to register for this training).

New project proposals can be submitted via the myDATA user portal. [myDATA](#) is your one-stop-shop for all things DataLab, from training to managing research projects. Once the project proposal is submitted it will take 1 to 2 weeks for ABS to conduct initial reviews. Once these reviews have been addressed, it will take 4 to 6 weeks to obtain the data custodian approvals.

## Charges

DataLab access incurs an [annual charge](#). This charge is based on the number of analysts with virtual machine access in a project.

Additional charges apply for customised [data integration services](#).

## Contact

Contact [data.services@abs.gov.au](mailto:data.services@abs.gov.au) for more information.



# Business Longitudinal Analysis Data Environment (BLADE)

BLADE is an economic data asset constructed by the Australian Bureau of Statistics (ABS) combining tax, trade, and intellectual property data with information from ABS surveys to provide a better understanding of the Australian economy and businesses performance over time.

This data asset can be used to understand how businesses fare over time and the factors that drive performance, innovation, job creation, competitiveness, and productivity. It can also be used to provide new insights into the development and evaluation of government policies, programs, and services.

## Datasets

The current BLADE asset contains data on all active businesses from 2001–02 to current, sourced from:

- ABS Business Register
- Australian Taxation Office: Business Activity Statements (BAS), Business Income Tax (BIT) filings and Pay as You Go (PAYG) summaries
- ABS surveys: Business Characteristics Survey (BCS), Management Capabilities Module (MCM), Economic Activity Survey (EAS), Business Expenditure on Research and Development (BERD), Private Non-Profit Expenditure on Research and Development (PNPERD), Survey of Research and Experimental Development, Government (GOVERD)
- Intellectual Property Australia: Intellectual Property Longitudinal Research Data (IPLORD)
- Department of Foreign Affairs and Trade Merchandise trade data

See [BLADE data inclusions and legislation](#) for a full list.

## Access, Charges and Contact

Same as PLIDA.

# ATO Longitudinal Information Files (ALife)

ALife is the tax and superannuation dataset constructed by the Australian Taxation Office (ATO). Approved researchers can access it in a safe and secure way.

## Datasets

The ALife is based on a random 10 per cent sample of tax returns that have been stripped of identifiers and subjected to a range of other techniques to preserve privacy. There are more than 400 variables. The data is available in 'csv' and 'dta' formats.

ALife includes a sample of:

- income tax return data since the 1999–00 financial year
- super data drawn from member contribution statements and self-managed super fund annual returns since 2007–08
- Higher Education Contribution Scheme (HECS), Higher Education Loan Program (HELP) and Student Financial Supplement Scheme (SFSS) data
- data collected from the ATO client register.

## Access

Approved researchers can access the ALife via Secure Unified Research Environment (SURE), a secure data laboratory hosted by the not-for-profit Sax Institute. Access to the ALife is limited to researchers from approved organisations. A head agreement must be in place before any of an organisation's researchers can apply for access. Approved researchers must commit to:

- safeguarding the data
- accessing and storing the data in a high-security data laboratory
- vetting all requests for data output.

Applicants typically require ethics approval for their research proposal. They also must explain how their work will benefit the community, such as by assisting the development or analysis of public policy. Researchers are only permitted to use the data for approved projects and in the precise way that has been approved. For more information, see [access guides for researchers](#) for more details.

## Charges

Sax Institute charges fees to cover the cost of accessing its secure environment. These fees are set independently of the ATO. The Sax Institute has published [a schedule of fees](#) to access the dataset.

## Contact

For inquiries, requests, or collaboration proposals, contact [alife@ato.gov.au](mailto:alife@ato.gov.au).

## Longitudinal datasets

This section lists 7 major longitudinal datasets that are funded by the Australian Government.

### Household, Income and Labour Dynamics in Australia (HILDA) survey

[HILDA Survey](#) is a household-based panel study that collects valuable information about economic and personal well-being, labour market dynamics and family life. The Survey is started in 2001 and follows the lives of more than 17,000 Australians each year. The HILDA Survey is funded by the Department of Social Services (DSS). The [Melbourne Institute: Applied Economic & Social Research](#) is responsible for the design and management of the Survey.

### Growing up in Australia: The Longitudinal Study of Australian Children (LSAC)

[LSAC](#) is a major study following the development of 10,000 children and families across Australia. It tracks children's development and life course trajectories in today's economic, social, and political environment. Information is collected on parenting, family relationships, education, childcare and health.

LSAC commenced in 2004 with 2 cohorts of 5,000 children each, aged 4 to 5 and 0 to 1 years. Participants are a representative sample of children of these ages across Australia at the time. Data is collected every 2 years. Parents (both residents and non-residents), carers and teachers of the children take part in the study. The children participate in the study when they reach an appropriate age.

LSAC is conducted in partnership between the DSS, the Australian Institute of Family Studies (AIFS) and Roy Morgan with advice provided by a consortium of leading researchers.

The study links to administrative databases, such as Medicare, Centrelink, Australian Early Development Census (AEDC), National Assessment Program – Literacy and Numeracy (NAPLAN), ABS Census of Population and Housing and Australian Curriculum, Assessment and Reporting Authority's (ACARA) My School data.

### Footprints in Time: The Longitudinal Study of Indigenous Children (LSIC)

[LSIC](#) is a study of Aboriginal and Torres Strait Islander children selected from different locations across Australia. It follows the development of around 1,700 Indigenous children and their families. Information is collected on parenting, family relationships, education, child and parent health, culture and community.

LSIC is funded by the Australian Government, led by an Indigenous majority Steering Committee, and managed by the DSS.

The study started in 2008. It involves 2 groups of Aboriginal and/or Torres Strait Islander children who were aged 0 to 24 months and 3.5 to 5 years in 2008.

Data is linked with the Australian Early Development Census (AEDC), the National Assessment Program – Literacy and Numeracy (NAPLAN) and Australian Curriculum, Assessment and Reporting Authority's (ACARA) My School data.

## Building a New Life in Australia: The Longitudinal Study of Humanitarian Migrants (BNLA)

**BNLA** is a study that follows close to 2,400 individuals and families to trace the settlement journey of humanitarian migrants from their arrival in Australia through to their eligibility for citizenship. Most of the participants live in Melbourne and Sydney and a small proportion live in regional or rural areas. They come from 35 countries and are aged from 16 to 80+. Almost all arrived in Australia in 2013, and experienced trauma and persecution before their arrival.

The study is conducted in partnership with DSS and AIFS. The study ran from 2013 to 2018 and recommenced in 2022.

Information is collected on housing, language proficiency, education, employment and income, health, and life satisfaction.

## Ten to Men: The Australian Longitudinal Study on Male Health

**Ten to Men** is a longitudinal study that follows participants over time and provide an important resource for improving the health of Australian men and boys. The study is being conducted by researchers at the AIFS and is funded by the Department of Health and Aged Care.

The study started in 2013 and follows approximately 16,000 boys and men aged 10 to 55 across the country through surveys and interviews.

The study is collecting information on wellbeing and mental health, use of health services, health-related behaviours, health status, health knowledge and social determinants. Ten to Men data are also linked with external datasets, including the Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS).

To apply to access *Ten to Men* data fill out the [Data Access Request Form](#) and send it to [ttmdatamanager@aifs.gov.au](mailto:ttmdatamanager@aifs.gov.au).

## Longitudinal Survey of Australian Youth (LSAY)

The **LSAY** provides information about young people and their pathways and is managed and funded by the Department of Education, with support from state and territory governments. It tracks young people as they move from school into further study, work, and other destinations. LSAY began in 1995, followed by additional cohorts in 1998, 2003, 2006 and 2009. The most recent cohort began in 2015. Nationally representative samples of over 10,000 young people start out in each cohort. Since 2003, participants have been recruited from Australian schools that take part in the Programme for International Student Assessment (PISA).

The survey participants enter the study when they are about 15 years old and are contacted once a year until they are 25. Information is collected on a wide range of school and post-school topics, for example, student achievement and aspirations, attitudes to school, social background, vocational and further education, employment and job seeking, satisfaction with various aspects of life.

LSAY records for the 2015 cohort have also been linked to several data sources including Australian Curriculum, Assessment and Reporting Authority's (ACARA) My School, National Assessment Program – Literacy and Numeracy (NAPLAN), Senior secondary administrative data, the National VET Provider Collection and the Higher Education Statistics Collection. Visit their [data linkage](#) page for more information.

LSAY unit record (including linked data) files are deposited with the Australian Data Archive (ADA) via [LSAY Dataverse](#). [Access to the data](#) is free via a formal request and registration process managed by the ADA.

## Longitudinal Literacy and Numeracy in Australia (LLANIA) dataset

The [Australian Education Research Organisation](#) (AERO) has created the first longitudinally linked dataset of Australian students' National Assessment Program – Literacy and Numeracy (NAPLAN) participation and results. This [LLANIA](#) dataset provides the ability to track students' performance and growth throughout their schooling and assess the effectiveness of learning interventions.

The linked dataset comprised data from students who were enrolled in the education system from 2008 to 2021 and holds data from more than 6 million students in total, with 25 per cent of them having fully linked data from Year 3 to Year 9. Students in Year 3 in 2009 to 2013 and 2015 were the cohorts that had the potential to reliably complete all 4 rounds of NAPLAN assessment, with fully matched records formed for 81 per cent to 86 per cent across calendar years.

# Australian Institute of Family Studies (AIFS)

AIFS conducts research into the wellbeing of Australian families, to inform government policy and promote evidence-based practice in the family services sector. AIFS runs numerous surveys on topics relevant to families, including sensitive topics such as violence and safety, mental health, gambling, and substance abuse. AIFS is an Accredited Integrating Authority and Accredited Data Service Provider (ADSP) for major longitudinal studies which conduct research into family, children, migrants, and men's health. They also provide secure and ethical access to a range of administrative and survey data sources for researchers, policy makers and service providers.

## Datasets

See a full list of [projects](#) for which AIFS provided data linkage services. AIFS is responsible for several of the longitudinal studies listed above, for example, Growing up in Australia, Building a New Life in Australia, and Ten to Men.

## Access

Access to the datasets of current AIFS projects requires either an individual or organisational licence. To apply, go to [National Centre for Longitudinal Data Dataverse](#).

Datasets derived from past research undertaken by AIFS are available from the [Australian Data Archive](#).

For data integration services, you need to submit a [request form](#) that outlines the research project, data needs and expected outputs. AIFS will review the request and contact you to discuss the feasibility, costs, and timelines of your project, and advise you on data access and governance arrangements.

Linkage and integration projects often involve the use of federal and/or state and territory data held by external data custodians or data owners. The relevant data custodian or owner must authorise all requests for access to data for linking purposes. As the Accredited Integrating Authority and ADSP, AIFS can handle these data access negotiations on your behalf.

The linked data is available through a Virtual Data Lab (VDL) service – a secure online environment where researchers can access and analyse de-identified data from various sources. The VDL provides a range of tools and software for data analysis and visualisation, as well as technical support and guidance from the data experts.

## Charges

AIFS charge for their data linkage and integration services based on the complexity and scope of the project, the data sources and variables involved, the data processing and analysis required, and the data access and delivery options preferred by the data users. AIFS provide an indicative cost for each project at the initial stage of the request process, and a final cost agreement after the project details and specifications are confirmed.

## Contact

To submit data linkage requests email request forms to [dliia@aifs.gov.au](mailto:dliia@aifs.gov.au).

# The Australian Data Archive (ADA)

ADA provides a national service for the collection and preservation of digital research data. ADA disseminates this data for secondary analysis by academic researchers and other users. The archive is based in the Centre for Social Research and Methods (CSRM) at the Australian National University.

## Datasets

ADA holds over 6,000 datasets from more than 1,500 projects and studies from 1838 through until the present day. Some of the most [popular datasets](#) are identifiable by methodology (for example, longitudinal) and by theme (for example, social attitudes surveys and health studies). You can also execute a search using the [Dataverse search tool](#).

## Access

ADA collection is managed through the Dataverse platform developed by the Institute for Quantitative Social Science at Harvard University. Data from the ADA is made available under 4 broad categories of access. The choice of access category is made by the depositor of the data:

- Open access
- General access
- Restricted access
- Special access

Access to data at each of these levels has an associated [user undertaking](#) that you must complete in order to be granted access.

## Charges

There is no cost for datasets unless specified in special conditions and restrictions. Contact ADA for more details.

## Contact

Email [ada@ada.edu.au](mailto:ada@ada.edu.au) to know more about ADA datasets.

# Australian Institute of Health and Welfare (AIHW)

The [AIHW](#) is an independent Australian Government agency working with health and welfare data. They collect and manage data on health and welfare issues, including from state, territory, and federal government agencies. They also make data available to state, territory and federal government agencies, universities, research centres, and non-government organisations. AIHW is also the national data linkage unit of the Population Health Research Network (PHRN).

## Datasets

AIHW holds more than 150 datasets covering fields as diverse as housing assistance, homelessness, perinatal health, disability, cancer, hospitals and hospital activity, alcohol and other drugs and mortality. See the [data collections](#) for a full list of available datasets. AIHW also holds and arranges access to the several [linked data assets](#) listed below.

### National Health Data Hub (NHDH)

The [NHDH](#), formerly the National Integrated Health Services Information (NIHSI) analytical asset, is a major national health data linkage system for health research and analysis. It comprises enduring health data from state and territory, and Commonwealth data sources. It brings together data on admitted patient care services (in public and private hospitals), emergency department services and outpatient services in public hospitals for all states and territories with the exception of Western Australia and the Northern Territory. The NHDH also includes Commonwealth data from the Medicare Benefits Schedule (MBS), Pharmaceutical Benefits Scheme (PBS), Repatriation Pharmaceutical Benefits Scheme (RPBS), Aged care services data, National Death Index (NDI) and the Australian Immunisation Register (AIR) data.

### National Disability Data asset (NDDA)

The [NDDA](#) will link data from Australian, state and territory government sources to better understand the life experiences of people with disability. Linking data from multiple sources can provide new insights into how people with disability interact across multiple support services and programs. This information will help improve programs and services for people with disability.

### National Aged Care Asset (NACDA)

The [NACDA](#) aims to bring together person-level data collected across aged care, health, and community service settings for aged care research purposes. The purpose of the NACDA is to better understand the pathways, experiences and outcomes of people receiving aged care, and the operation and performance of the aged care system. The first iteration of the National Aged Care Data Asset consists of linked data from National Aged Care Data Clearinghouse (NACDC), Medicare Benefits Schedule (MBS), Pharmaceutical Benefits Scheme (PBS) and Repatriation Pharmaceutical Benefits Scheme (RPBS), National Death Index (NDI), Hospital collections (admitted patient, non-admitted patient, emergency) and Australian Immunisation Register (AIR).

### Child Wellbeing Data Asset (CWDA)

The [CWDA](#) aims to improve outcomes for children, young people and their families. It will enable researchers to access enduring integrated data safely, securely, and ethically across sectors to analyse



children and young people's wellbeing and pathways through government services over time. The CWDA is in development and planned to be ready for use in early 2025. It will include Australian Early Development Census (AEDC), Child Protection National Minimum Data Set (CP NMDS), Youth Justice National Minimum Data Set (YJ NMDS), National Community Mental Health Care Database (NCMHCD), Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS).

## COVID-19 register and linked data asset

The [COVID-19 register and linked data asset](#) combines COVID-19 case data with a range of other datasets, including hospitals and deaths data. Linking COVID-19-related datasets provides new insights into the health outcomes for people who have been diagnosed with the disease, and the effect that COVID-19 has had on the health system and broader community. The linked dataset offers researchers the opportunity to explore a range of issues associated with the pandemic.

## Access

Some data collections cannot be made publicly available. Access to others need special approval, which in some cases must be given by the data providers. Where data is not publicly available, you may [request data](#) via the data on request service which provides clients with custom data. Learn more about the data access process and requirements by the AIHW through their [corporate brochure](#).

The Department of Social Services (DSS) also makes Data Over Multiple Individual Occurrences (DOMINO) accessible through the AIHW.

AIHW is an Accredited Data Service Provider (ADSP) and undertake [data linkage](#) to bring together data from multiple sources to better understand the health and welfare of Australians and the interactions they have with health and welfare services. Their Data Integration Services Centre (DISC) assists researchers with customised data linkage projects that will have a clear benefit to the community. AIHW provide a range of data linkage service from the beginning of the project until completion. Their services include advice on project design and methodology, ethics approvals, negotiating access to relevant datasets, file transfer, access to data through secure environment and safely destroying data.

## Charges

Cost recovery applies for AIHW data linkage services. [Costs](#) will vary, depending on the size and complexity of the project. Contact the AIHW DISC via [linkage@aihw.gov.au](mailto:linkage@aihw.gov.au) to request a quote, or if you have any questions about data linkage services.

Charges may also apply to data access at instances where the data is available through tables prepared by the AIHW for specific requests.

## Contact

Email [info@aihw.gov.au](mailto:info@aihw.gov.au) for more information.

# Department of Social Services (DSS)

DSS funds or conducts longitudinal surveys in partnership with other organisations.

## Datasets

There are currently 4 active surveys detailed above (HILDA, LSAC, LSIC, and BNLA). The National Centre for Longitudinal Data (NCLD) is responsible for these studies.

## Access

The NCLD provides access to new releases of data for the 4 longitudinal studies through the Australian Data Archive (ADA) Dataverse platform. To apply for access to any of the DSS Longitudinal Studies datasets see [DSS Longitudinal Studies Dataverse](#).

Dataverse allows authorised data users to download the datasets to their password-protected PC, secure network, or approved cloud.

## Charges

Dataverse provides no cost data access to authorised users.

## Contact

Contact NCLD ([nclresearch@dss.gov.au](mailto:nclresearch@dss.gov.au)) and ADA ([ada@anu.edu.au](mailto:ada@anu.edu.au)) for more details.

# State and territory data agencies

## Population Health Research Network (PHRN)

PHRN is a national collaboration that enables existing data from around Australia to be brought together and made available for important research.

PHRN supports a network that enables approved population data from around the nation to be safely and securely analysed by researchers in a secure remote access data laboratory. [Secure Unified Research Environment](#) (SURE) is one of the remote-access data research laboratories purpose-built for the analysis of linked, routinely collected data. It allows researchers to remotely access and analyse the approved data extracts provided by data custodians for their research project. Researchers who are interested in applying for access to linked data for cross-jurisdictional and multi-jurisdictional research projects should apply via the PHRN [Online Application System](#).

PHRN comprises of a network of data linkage units nationally as well as in each Australian state/territory. AIHW is the national data linkage unit of PHRN. Details for each of the state/territory data linkage units is provided immediately below.

## NSW & ACT: Centre for Health Record Linkage (CHeReL)

CHeReL is dedicated to helping researchers, planners and policy makers access linked health data about people in the NSW and ACT. It provides advice on the design, cost, feasibility, and process of linkage studies. They offer different types of record linkage services based on the needs. See their [services](#) for more details.

### Datasets

CHeReL can link consenting participants to a broad range of routinely collected data from NSW Health and other administrative or research data collections. Examples of routinely collected NSW Health data and accompanying data dictionaries can be accessed at [datasets](#).

A Biospecimen Linked Standardised Data Asset (BLSDA) is being developed to simplify access to linked NSW Health data. The BLSDA will be a core minimum health data asset of linked NSW Health data collections for the participants in a number of strategic biospecimen collections held in the NSWHSB. The NSW Health data asset will be available off the shelf at no cost to eligible researchers.

### Access

If you are in NSW Health and will be using data for evaluation purposes you can [contact](#) them directly. For research related purposes, there are several steps involved in gaining approval which are detailed for [NSW](#) and [ACT](#), respectively.

### Charges

CHeReL charges a fee for the linkage of records. The [fee matrix](#) shows the estimated fees for their linkage services, depending on the cohort size and the number and types of datasets to be linked.

## Contact

For more information, contact the CHeReL at [MOH-CHeReL@health.nsw.gov.au](mailto:MOH-CHeReL@health.nsw.gov.au).

## VIC: Centre for Victorian Data Linkage (CVDL)

CVDL is a data linkage unit established in 2009 and provides a range of services to researchers and other Victorian Government departments. It is based in the [Victorian Agency of Health Information \(VAHI\)](#) in the Department of Health.

## Datasets

CVDL maintains the Victorian Linkage Map (VLM) and Integrated Data Resource (IDR), a collection of almost 40 [datasets](#) which are linked on a monthly basis.

## Access

CVDL's [online application form](#) is to be used for data linkage requests requiring access to Victorian datasets available from the CVDL. For projects requiring data from Victoria and one or more other jurisdictions, see [cross-jurisdictional projects](#). See the [data linkage assessment and approval process](#) for more information.

The data is accessed via Victorian data Access Linkage Trust (VALT), the CVDL secure data analysis environment. It includes standard statistical analysis software packages (with the exception of SAS), including R, Python and SQL. Stata can be included upon request; however, the user will need to have their own Stata licence.

## Charges

CVDL operate on a cost recovery basis and charge fees for data linkage projects and for the use of VALT virtual machines to access and analyse data. The cost for data linkage projects varies according to the complexity of the request, particularly where there are requests for data that aren't routinely linked in the IDR, if there are cohort files that need to be linked, whether the project involves multi-jurisdictional or cross-jurisdictional linkage, and whether additional legal review is required – for example, development of a privacy impact assessment. You can contact CVDL to discuss your project requirements and to obtain a formal quote of linkage costs.

## Contact

Contact CVDL via [cvdl@vahi.vic.gov.au](mailto:cvdl@vahi.vic.gov.au).

## QLD: Data Linkage Queensland (DLQ)

The Statistical Analysis and Linkage Unit within the Statistical Services Branch (SSB) of Queensland Health is also known as [Data Linkage Queensland](#). DLQ routinely links many Departmental and other health-related datasets, and performs data linkage to facilitate and promote health service planning, management, monitoring and evaluation, and research within Queensland. Additionally, DLQ is responsible for much of Queensland's health data linkage for Commonwealth requirements, national committees, and cross-jurisdictional research projects. As a member of the Population Health

Research Network (PHRN), DLQ works with data linkage units in other Australian jurisdictions to link health datasets from other states and territories for use in national research.

## Datasets

SSB is the custodian for some data collections. Detailed information about individual datasets is available in the list of [commonly requested data items and collection manuals](#). There are also some datasets that have been linked or those commonly requested for departmental or research linkage requests. See a list of these [linked datasets](#) for more details.

## Access

Requests for access to data from SSB's data collections are considered on a case-by-case basis. Refer to the [data request decision tree](#) to see which form of application is appropriate for your request.

To request linked data contact DLQ. Release of identifiable or potentially identifiable data requires data custodian approval. SSB also offers a data linkage service to researchers and Queensland Health staff for approved projects. For further information on data linkage in Queensland Health, see the [Queensland Data Linkage Framework](#).

## Charges

Contact DLQ to know more about the associated costs.

## Contact

Contact DLQ at [DLQ@health.qld.gov.au](mailto:DLQ@health.qld.gov.au) for any data linkage enquiries.

## SA & NT: SA NT DataLink

[SA NT Datalink](#) is the data linkage research infrastructure supporting population-based data linkage in South Australia and the Northern Territory and works closely with other jurisdictional data linkage units, including the Commonwealth. As a member of the Population Health Research Network (PHRN), it works in collaboration with other Australian linkage units, including the Commonwealth.

## Datasets

SA NT DataLink routinely updates data sources. See [datasets](#) for a complete list and the years for which the data is available. They also facilitate the linking of other data for one off approved projects. This includes project cohorts, registries, and other non-routinely updated data sources from both the Non-Government and Government sectors. See a list of examples for [linkage projects](#).

## Access

Approval to receive access and use of linked unit record level data through SA NT DataLink, is subject to data custodian approval, and ethical approvals as required, with deeds of confidentiality & compliance signed by people accessing data.

Researchers complete a PHRN [online application](#) to initiate the application and approval process for SA, NT data, or other Australian jurisdictional data. See the [application process](#) for more details.

Australian Government data at a unit record level including MBS, PBS, Centrelink, Aged Care, Disability (NDIS), National Death Index, National Immunisation Registry is securely stored and accessed remotely in a secure data laboratory.

## Charges

Due to limited funding, there is a partial cost-recovery fee charged for accessing linked data, based on the amount of work needed to be performed. See [Access and Pricing Policy](#) for more details.

## Contact

Any inquiries related to data linkage should be directed to [SANTDataLink@unisa.edu.au](mailto:SANTDataLink@unisa.edu.au).

## WA: Western Australia Data Linkage System (WADLS)

The [WADLS](#) was established as a collaboration between the Department of Health WA, the University of Western Australia (UWA), Telethon Kids Institute (then the Telethon Institute for Child Health Research) and Curtin University. The data linkage team manages the WADLS. They create, store, update, and extract links between population-based administrative and research data collections (both health and non-health related) in WA.

## Datasets

Some data collections are held within the WADLS, while others are held in the linked data warehouse. See the [data collections](#) for most frequently requested datasets and collections available via data linkage services.

## Access

[Datasets can be requested by submitting an Application](#), or through contacting the relevant data custodian. For more detailed information on a specific data collection, applicants are advised to contact the data custodian directly. See the [Application process](#) for more details.

The process of extracting data for projects involves multiple teams at the Department of Health, as well as the assistance of data collections from the Department of Health and other external agencies. Once necessary approvals have been granted (for example, by relevant ethics committees or stakeholders), the data outputs team extracts and delivers the requested data in a secure method.

## Charges

Charges for data linkage services are outlined in the [Prioritisation and Costing Framework](#). A cost estimate for the data linkage will be provided during the feasibility assessment of the application. Upon delivery of the data the final cost will be determined.

## Contact

You can contact [DataServ@health.wa.gov.au](mailto:DataServ@health.wa.gov.au) for research data related queries and [ISPDClientServices@health.wa.gov.au](mailto:ISPDClientServices@health.wa.gov.au) for non-research data related queries.

## TAS: Tasmanian Data Linkage Unit (TDLU)

TDLU brings together Tasmanian data collections and other approved data sources to create linked datasets. The data they link is sourced from multiple administrative and clinical datasets from various sectors including health, education, and welfare. TDLU undertakes both state and national data linkage projects.

### Datasets

Several datasets have been linked into the Master Linkage Map and are available for research, subject to relevant approvals. See [datasets available](#) for a full list.

### Access

Whether you are applying for a project in Tasmania, or a cross- jurisdictional or multi-jurisdictional project using Tasmanian datasets, they encourage you to discuss your project with TDLU staff before applying. There is step by step [application process](#) involved in order to gain access to the linked data. Researchers complete a PHRN [online application](#) to initiate the application and approval process. See helpful [resources](#) to prepare your application.

### Charges

There is cost involved for data linkage and it increases based on the number of datasets, time period of data required, and complexity of the project. The TDLU has 3 stages of invoicing. contact TDLU to obtain a quote. An indicative quote will be supplied by the TDLU after preliminary discussions, with a revised quote provided on completion of the design phase. Data custodians may charge a fee for extraction of data from their collections; the TDLU will advise if any additional fees are likely.

### Contact

You can contact the TDLU at [menzies.tdlu@utas.edu.au](mailto:menzies.tdlu@utas.edu.au).

## NSW Bureau of Crime Statistics and Research (BOCSAR)

BOCSAR is a statistical and research agency within the Department of Communities and Justice. One of their roles is developing and maintaining statistical databases on crime and criminal justice in NSW.

### Datasets

BOCSAR has compiled a collection of sizeable crime, court, and custody datasets, most are updated annually or quarterly. They are intended for use by people who want to perform their own analysis. The datasets included are, for example, criminal offences by geography, short-term and long-term crime trends, special offence data, drink spiking incidents in NSW. See their [Open data](#) page for a complete list of datasets.

Unit record level criminal incident and reoffending data are also available with restricted access.

## Access

Some datasets are available with open access. Customised datasets are also available from BOCSAR on request. To access unit level criminal incident and reoffending data you will have to meet specific conditions, including an ethics approval for the proposed research. Refer to the [Unit Record Reoffending Data](#) and [Unit Record Criminal Incident Data](#) guidelines for further details.

## Charges

If the dataset you are after is relatively straightforward for BOCSAR staff to, it will be free of charge and will be available within 2 working days. If you are after data that is more complex to, it will incur a [cost](#) and will take up to 10 working days to prepare.

## Contact

Email [bcsr@dcj.nsw.gov.au](mailto:bcsr@dcj.nsw.gov.au) for more details.