



Using PLIDA & Linked Administrative Data to Advance Evidence on Equity in Work, Markets and Society.

Ed Low-Thurley

Director, Data Integration Assembly

AUSTRALIAN BUREAU OF STATISTICS
Informing Australia's important decisions

What is Data Integration?

- ▶ In the ABS, data integration is the process of combining data from two or more sources at the unit level (person, business, location) in a safe, secure environment to provide new data insights for statistical and research purposes.
- ▶ We are an accredited data service provider. ABS has the infrastructure and capability to undertake high risk data integration projects for statistical and research purposes. We don't do compliance and enforcement work. ABS has a responsibility to protect and use data for the benefit of the entire community.
- ▶ ABS works as a trusted partner with government, research institutions and business on data integration projects that have a clear public benefit. The data we integrate comes from the Census and surveys we run, and from other government agencies administrative effort.

Connecting admin data for Australia's future

Why do we link and share data?

1. To produce more timely, relevant, and granular insights for public decision-making

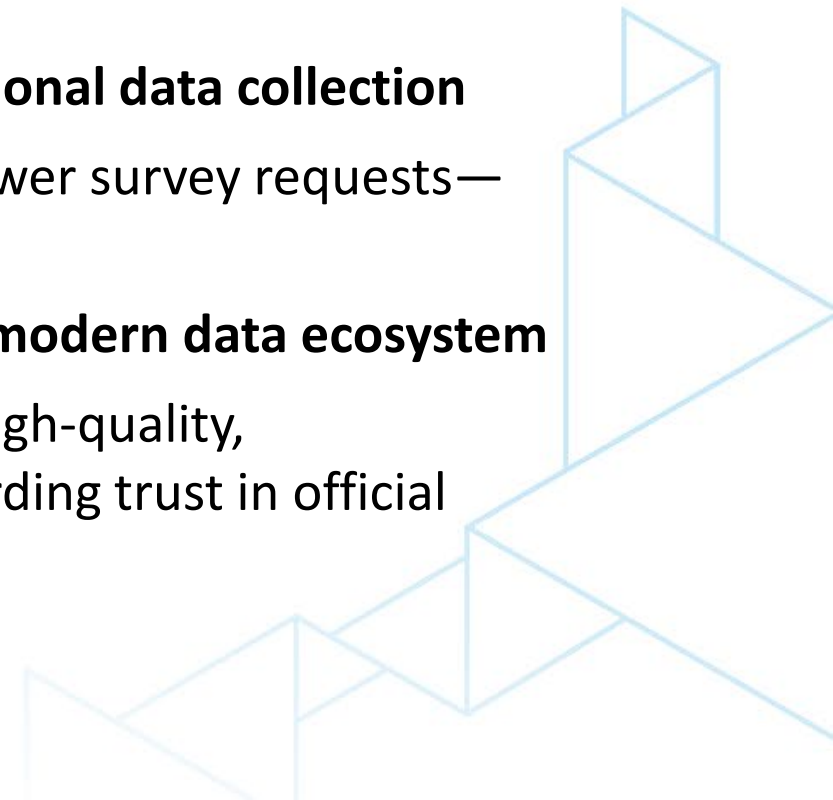
Integrated administrative data gives policymakers the evidence they need to design effective interventions in areas such as health, education, employment, business productivity, and demographic change.

2. To increase efficiency and reduce the cost and burden of traditional data collection

Governments save resources, while citizens and businesses face fewer survey requests—creating a more streamlined national data system.

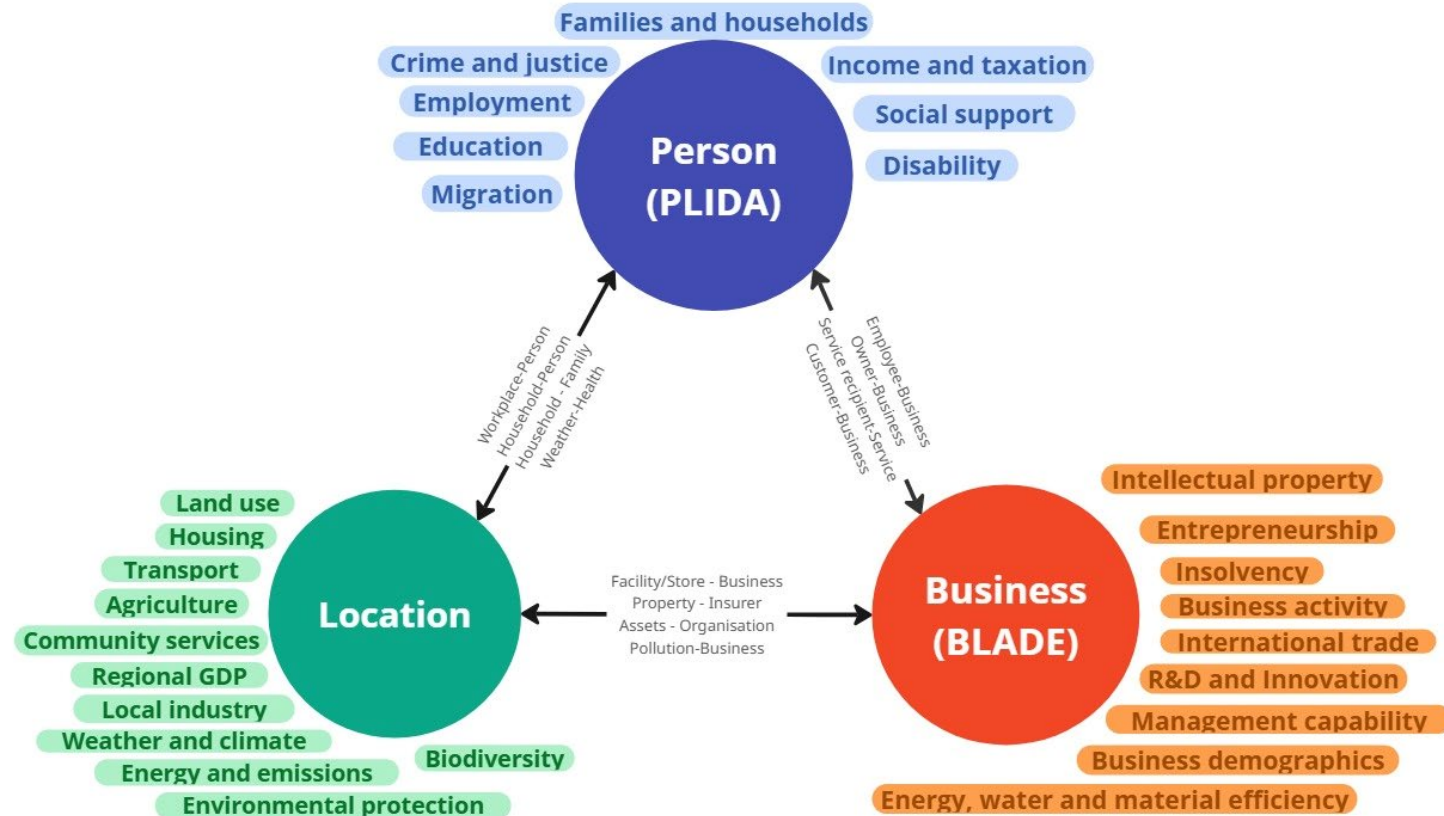
3. To strengthen national data governance and build a coherent, modern data ecosystem

Data integration strengthens national data governance, ensuring high-quality, privacy-protected information is available to society while safeguarding trust in official statistics.



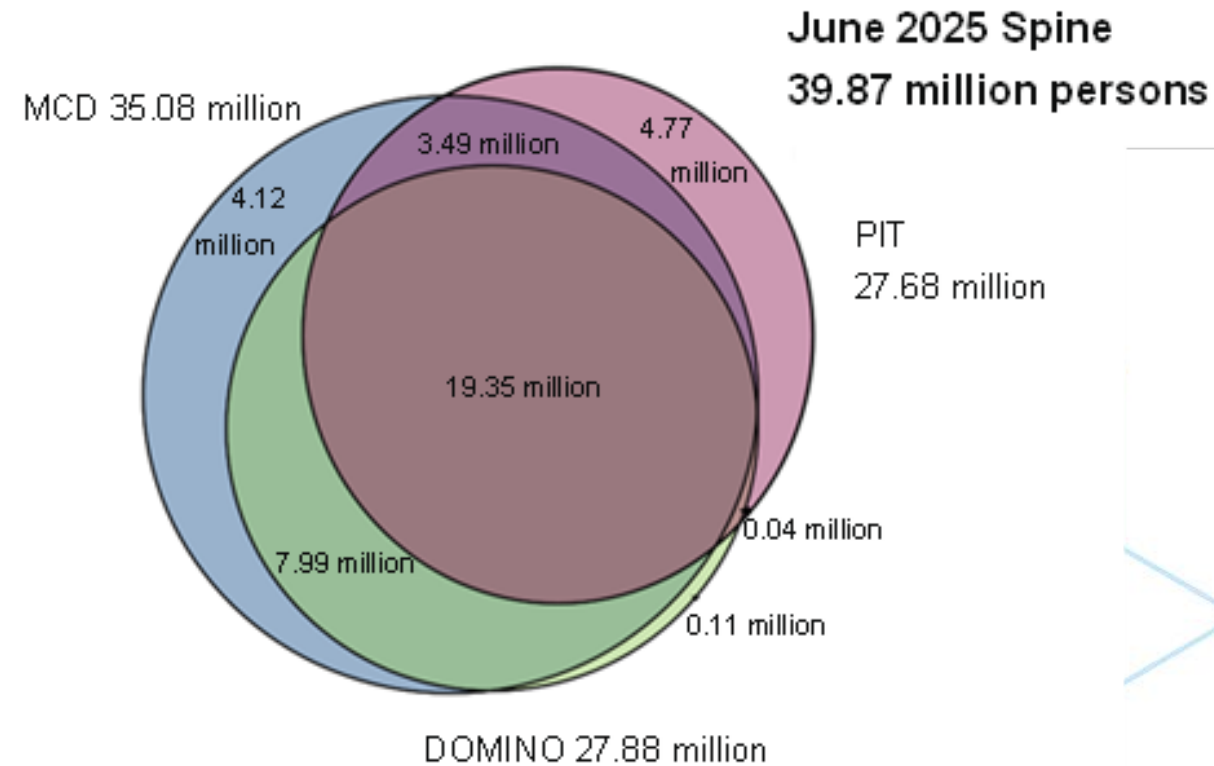
An integrated view of Australia

- ABS integrated data consists of **datasets from a wide range of organisations** (data custodians) around Australia.
- **Each dataset is converted into a “linkable” module.** Each module includes information that can be linked at the person-, business, or location-level, and sometimes all three.
- This **supports multidisciplinary research and statistics** on social, health, economic and environment issues.
- There are **216 modules** that have been created.
- Of these 216,
 - 121 are person modules, 91 business modules and 4 location modules.
 - 79 are available for re-use on an enduring basis (37 PLIDA and 42 BLADE)
 - 137 are either in development or have some restriction or limited re-use.
- New areas of development in wealth, housing, environment, veterans, health, criminal-justice, disability.



Person Linkage Spine

- The ABS Person Linkage Spine is foundational data linking infrastructure that **underpins the PLIDA data asset**
- The Spine is constructed from **three high-quality data sources**:
 - Medicare Consumer Directory (MCD)
 - DOMINO Centrelink Administrative Data
 - Personal Income Tax (PIT)
- The Spine includes data **from January 2006** through June 2025 and aims to represent all persons who were resident in Australia at any point during this period
- **The Spine comprises 39.87 million unique persons**
- It is expected that the next spine update will be delivered in late 2026 and will extend the spine reference period to **30 June 2026**

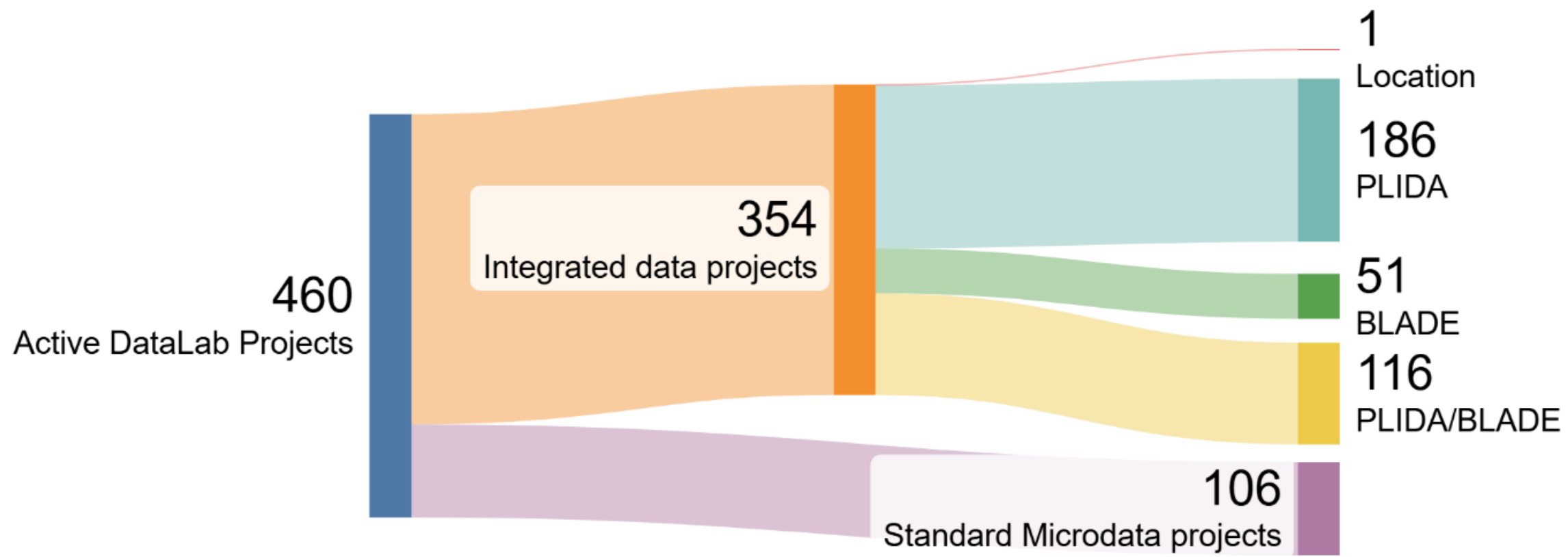


- The spine (v8) will be upgraded with births, deaths and migration data for the 2026 spine build



Projects accessing integrated data – Feb 2026

At 27 February 2026, there are **460** active projects in the DataLab accessing integrated data (**354**), and standard microdata (**106**). Of the **354** projects currently accessing integrated data, **1** uses LOCATION data, **186** use PLIDA, **51** use BLADE and **116** use PLIDA and BLADE.



Five Safes Framework

 Safe people

 Safe projects

 Safe settings

 Safe data

 Safe output

Is the researcher authorised to access and use the data appropriately?

Is the data to be used for an appropriate purpose?

Does the access environment limit unauthorised use?

Has appropriate and sufficient protection been applied to the data?

Are the statistical results non-disclosive?

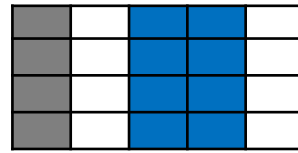
Managerial
controls

Statistical
controls



PLIDA enhancements – Core modules

Source
datasets in
PLIDA



- Core Demographics
- Core Locations
- Core Relationships
- Core Scoping
- Core Income

+

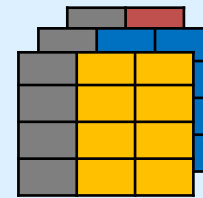
- **Core Disability Flags**
- **Housing and Wealth**
- **Education**
- **Indigenous**
- **Health**

Core Module

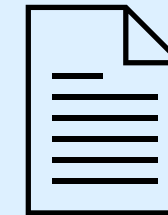
=



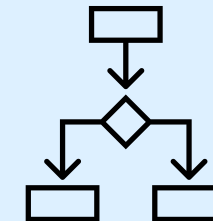
+



+



+



1. Combined table(s)

2. Single source
tables

3. Explanatory
notes

4. Pseudocode
for constructing
1 from 2

Visibility of LGBTQ populations in PLIDA is dependent on what's captured in the source datasets – but it is increasing!

2021 Census counted 78,425 same-sex couples living together in Australia.

► Core Demographics

- Sex or gender

Sex and gender often conflated in source data. Current approach not perfect, but it's transparent and a start.

- 1 = consistently recorded 'M'
- 2 = consistently recorded 'F'
- 3 = inconsistent or 'other' gender recorded

► Core Locations

- Where people live (de-identified)

Allows researchers to associate who people are living with.

► Core Relationships

- Spouse & parent-child relationships

- Combined with Core Demographics, can identify same-sex relationships (limited to 2 people).
- Building block for family structures.

2026 Census – new data items on sexual orientation and gender.

Challenges & opportunities

- ▶ Sex and gender is often conflated across different administrative data sources
 - e.g. Medicare gender markers may be viewed as sex assigned at birth
 - Inconsistent updating of 'gender markers' in data sources
 - Legacy IT and data systems
- ▶ Sexual orientation is rarely directly captured, often requiring inferences (e.g. same-sex spouse reported in Tax Returns).
- ▶ Lags in administrative data.
- ▶ This means only certain subsets of LGBTQ populations are visible (e.g. people engaged with healthcare or who reported living in a same-sex relationship).
- ▶ Questions on sexual orientation and gender identity in upcoming Census will provide increased opportunity to investigate population level effects.

Despite these challenges, PLIDA is already being used for research for and by LGBTQ populations.

- ▶ Cross sex hormone therapy use in PBS records 2012-2024 → >30,000 trans people on gender-affirming hormone therapy (Saxby et al., [2026](#))
- ▶ All cause and cause specific mortality risk (Saxby et al., [2026](#))
- ▶ Mental healthcare use and healthcare costs following initiation of gender affirming care (Saxby et al., [2026](#); Saxby & Nolan [2026](#))
- ▶ How community attitudes shape healthcare use and long-term health outcomes among individuals in same-sex relationships (Saxby et al., [2020](#); [2024](#))
- ▶ National HIV Data Linkage Study (UNSW)
 - Monitoring HIV pre-exposure prophylaxis (PrEP) uptake in PBS records (Chan et al., [2026](#))
- ▶ Mental health outcomes among trans Australians in National Study Mental Health and Wellbeing (Ma et al., [2025](#))

Access outside of Australia

- ▶ ABS now supports DataLab access outside Australia as an ongoing and cost-recovered service.
- ▶ This initiative supports international collaboration and enables secure access to Integrated data by overseas researchers and organisations.
- ▶ This service and the high-level process is outlined on the ABS Data Integration Service webpage: [Access outside of Australia](#)

The ABS has formalised access to the DataLab for researchers located outside Australia.



Providing this service means the world's best minds can bring their skills and experience to bear on Australian policy challenges, and enables Australian based researchers to continue their important work when travelling overseas.

Access outside of Australia: Scenarios

Scenario	Description and summary of assessment / approval requirements
Scenario 1: New overseas <u>project</u>	A project led by or involving overseas entities. Assessment: Standard safe project checks plus enhanced scrutiny for public interest value. Approval: Data Custodian approval
Scenario 2: Overseas <u>entity</u> joining existing Australian-based <u>project</u>	An overseas organisation added to an existing Australian-based project. Assessment: Lead organisation provides justification for collaboration benefits. Approval: Data Custodian approval
Scenario 3: Offshore Australian researcher accessing existing DataLab project	Australian-affiliated researcher accessing DataLab from outside Australia for a defined period. Assessment: Requires justification for continued access, endorsed by a Responsible Officer. Approval: Delegated to the Australian Statistician (person-level approval).
Scenario 4: Overseas researcher joining existing DataLab project	Researcher affiliated with an overseas entity applying for DataLab access. Assessment: ABS governance, security and legal teams review, alignment with ABS requirements confirmed. Approval: Delegated to the Australian Statistician.
Scenario 5: Overseas <u>entity</u> joining existing overseas project	Adding an overseas organisation to an existing overseas project. Assessment: ABS reviews legal context and undertakings. Approval: Delegated to General Manager, Statistical Production and Digital Services.

Questions?

