



Australian Bureau of Statistics

Privacy Impact Assessment Update

2024-25 Person Level Integrated Data Asset

Date of analysis– 9 December 2024

Date of finalisation– 11 April 2025

Contents

Contents	2
Part A Executive Summary	4
1. Introduction 4	
2. Overview of the 2024-25 PLIDA Update	5
3. This PIA Update process	6
4. Structure of this PIA Update report.....	6
5. Summary of findings	7
6. Recommendations	8
Recommendation 1 NFP organisations as Data Custodians	8
Recommendation 2 Approach to determining ‘inherently sensitive survey data’	9
Recommendation 3 Justifying combination of the ‘Linker’ and ‘Librarian’ roles within the ABS	10
Recommendation 4 Use of ANDII ICT solution for PLIDA processing and to ‘re-use’ PLIDA data for the NDDA	10
Recommendation 5 Governance arrangements for the Core Modules and the LCDI.....	11
Recommendation 6 Additional measures for approved research projects to support surveys	12
Part B Methodology and Assumptions	13
7. Our methodology	13
8. Assumptions and qualifications.....	15
Part C Overview of PLIDA	16
9. PLIDA Governance.....	16
10. Benefits of PLIDA.....	16
11. PLIDA data	17
Part D Summary of Privacy Compliance	19
12. Summary of Compliance	19
Part E Privacy Analysis - Inclusion of new and expanded data types in PLIDA.....	21
13. Introduction	21
14. PIA Topic 1: Linking high level administrative crime and justice data to PLIDA	21
15. PIA Topic 2: Linking ‘not-for-profit’ (NFP) organisational data to PLIDA	26
16. PIA Topic 3: Inclusion of ‘sensitive’ and ‘inherently sensitive’ survey data in PLIDA .	29
Part F Privacy Analysis - Implementation of new or changed data handling practices for PLIDA.....	31
17. Introduction 31	
18. PIA Topic 4: Combining the ‘Librarian’ and ‘Linker’ roles.....	31
19. PIA Topic 5: Migrating SDIE into ANDII ICT environment for data integration purposes	33

20.	PIA Topic 6: Reusing PLIDA Data for the NDDA/ANDII.....	35
21.	PIA Topic 7: International researcher access to PLIDA data in the DataLab	37
Part G	Privacy Analysis – Expanded outputs from PLIDA	40
22.	Introduction	40
23.	PIA Topic 8: Creation of new PLIDA Core Modules and the Life Course Data Initiative (LCDI)	40
24.	PIA Topic 9: Expanding the use of PLIDA to support statistical production (including the Census and ABS Household Surveys).....	46
25.	PIA Topic 10: Publication of aggregated PLIDA data.....	48
Part H	Glossary	50
Attachment 1	Materials Reviewed	54
Attachment 2	PLIDA Information Flow Diagram – Current Data Handling Practice	55
Attachment 3	PLIDA Information Flow Diagram – Proposed Data Handling Practice....	56
Attachment 4	PLIDA datasets.....	57

Part A Executive Summary

1. Introduction

- 1.1 The Person Level Integrated Data Asset (**PLIDA**), formerly known as the Multi-Agency Data Integration Project (**MADIP**), is a secure, enduring, person-based research data asset that combines broad sets of information about Australian populations. PLIDA creates a comprehensive picture of Australian residents over time and facilitates the use and re-use of public data for statistical analysis and research purposes.
- 1.2 The Australian Bureau of Statistics (**ABS**) is the accredited integrating authority for PLIDA and is responsible for combining the datasets in PLIDA, for providing access to PLIDA data to authorised users for approved research projects, for ensuring the security of data contained within PLIDA, and ensuring that all research results and other outputs from the use of PLIDA data are produced in a manner that is unlikely to enable the identification of a particular person.
- 1.3 PLIDA has evolved over time, and the ABS has undertaken three major privacy impact assessments (**PIAs**) in relation to PLIDA to date (an initial PIA in 2018 by Galexia (**the 2018 PIA**), a PIA update undertaken by the ABS in 2019 (**the 2019 PIA Update**), a further PIA update undertaken by Maddocks in 2022 (**the 2022 PIA Update**). Maddocks also undertook a PIA on the Expanded Health Data Linkage to PLIDA in 2024 (**the Expanded Health Data PIA**)). Each subsequent PIA process considered actual or planned changes or updates to PLIDA since the previous PIA process.
- 1.4 The ABS has also conducted various project-specific PIAs involving PLIDA data during this time, including detailed consideration of the linkages of new datasets to PLIDA on a case-by-case basis (for example, a PIA process was conducted in November 2020 for the linkage of data from the Cancer Institute of NSW).¹
- 1.5 The ABS has engaged Maddocks to conduct a further PIA update process in 2024-25, to:
 - 1.1.1 consolidate analysis of the privacy impacts of changes or updates to PLIDA since 2022 (taking into account the previous PIA processes and other privacy consideration that has already been given to address those changes or updates);
 - 1.1.2 consider potential changes to PLIDA governance, to ensure continued compliance with the Privacy Act, the Australian Privacy Principles (**APPs**), and privacy best practice; and
- 1.5.1 facilitate the ABS continuing to take a 'privacy by design' approach to PLIDA.

¹ Copies of the relevant PIAs are available at: <https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-impact-assessments>.

2. Overview of the 2024-25 PLIDA Update

2.1 This 2024-25 PIA Update (**PIA Update**) covers the following topics:

- 2.1.1 the inclusion of new or expanded data in PLIDA (in terms of type and volume), where that data will be prepared for linkage in accordance with existing PLIDA processes – the specific issues are:
- (a) **PIA Topic 1:** new variables drawn from datasets received from government entities, which reflect high-level information about data subjects' interactions with the justice system;
 - (b) **PIA Topic 2:** data received from entities who operate in a 'not for profit' manner (**NFP organisations**); and
 - (c) **PIA Topic 3:** consideration of the existing privacy practices for the handling of data which was originally collected by entities through surveys;
- 2.1.2 potential implementation of new or changed data handling practices for PLIDA, specifically:
- (a) **PIA Topic 4:** the potential combining of two internal ABS roles (known as the 'Linker' and 'Librarian' roles) for the handling of identified data received from the entities which provide that data to the ABS for PLIDA (**Data Custodians**);
 - (b) **PIA Topic 5:** changing the ICT system that is currently used to prepare PLIDA data, to another ICT system which has been built and is controlled by the ABS (that is, the Australian National Data Integration Infrastructure (**ANDII**)) ICT solution will be used for PLIDA data integration purposes, rather than the Secure Data Integration Environment (**SDIE**));
 - (c) **PIA Topic 6:** consideration of the mechanisms that have been implemented to 'reuse' PLIDA data if the relevant Data Custodian has approved the use of that data for the National Disability Data Asset (**NDDA**), in addition to its use for PLIDA; and
 - (d) **PIA Topic 7:** the potential access to analytical PLIDA data in the DataLab by authorised researchers from outside of Australia; and
- 2.1.3 expanded outputs from PLIDA data, including:
- (a) **PIA Topic 8:** the creation of new enduring modules which can be used by approved researchers (**Authorised Users**) for approved PLIDA projects. In conjunction with this PIA Topic, we also considered the implementation of the Life Course Data Initiative (**LCDI**);
 - (b) **PIA Topic 9:** use of data in PLIDA to supplement or replace data items collected via ABS surveys (including the Census) and to undertake more efficient survey operations; and
 - (c) **PIA Topic 10:** the potential creation of outputs containing integrated and aggregated analytical PLIDA datasets for publication on the ABS website, subject to the usual PLIDA and other ABS processes for the release of data.

3. This PIA Update process

- 3.1 Undertaking a PIA process is consistent with the requirements of the Privacy (Australian Government Agencies – Governance) APP Code 2017 (**APP Code**), which has applied since 1 July 2018. The APP Code requires agencies to undertake a written PIA for all ‘high privacy risk’ projects or initiatives that involve new or changed ways of handling personal information, but undertaking a PIA for other projects which have not been assessed as representing ‘high privacy risk’ is also consistent with privacy best-practice.
- 3.2 Undertaking subsequent supplementary processes, on a regular basis, to re-examine the potential privacy impacts of an APP entity’s undertaking as plans and circumstances change, is consistent with privacy best practice in Australia. We commend the ABS for undertaking subsequent PIA processes following the completion of the original 2018 MADIP PIA process for PLIDA, which has ensured that potential privacy impacts of specific projects or changes to PLIDA were subject to separately documented privacy threshold assessments, PIAs or other privacy advice.
- 3.3 This PIA Update will be a further measure consistent with the APP Code and privacy best practice. It is intended to help the ABS and the PLIDA Board manage any new or increased privacy risks and impacts, and may also assist in informing stakeholders about the on-going importance of privacy considerations being a critical consideration for PLIDA.

4. Structure of this PIA Update report

- 4.1 This PIA Update report consists of:
- 4.1.1 this **Part A**, which includes:
- (a) a description of the 10 topics covered by this PIA Update (see paragraph 2.1);
 - (b) a summary of our findings in relation to those 10 topics (see paragraph 5); and
 - (c) our recommendations to address identified privacy risks (see paragraph 6);
- 4.1.2 **Part B**, which sets out details of our methodology and further information about the scope of our work;
- 4.1.3 **Part C**, which contains an overview of PLIDA, to supplement the introduction in this **Part A**);
- 4.1.4 **Part D**, which contains a table setting out our overall compliance risk ratings against each APP in relation to all of the topics covered by this PIA Update report; and
- 4.1.5 **Part E**, in which we discuss specific privacy impacts, risks and issues associated with topics about the inclusion of new and expanded data types in PLIDA (PIA Topics 1-3);
- 4.1.6 **Part F**, in which we discuss specific privacy impacts, risks and issues associated with topics about the new or changed data handling for PLIDA (PIA Topics 4-7);
- 4.1.7 **Part G**, in which we discuss specific privacy impacts, risks and issues associated with topics about expanded outputs in PLIDA (PIA Topics 8-10); and
- 4.1.8 **Part H**, which contains a glossary of defined terms.

5. Summary of findings

- 5.1 PLIDA has developed over time to include robust governance frameworks and arrangements to protect the privacy of data subjects. In our view, continued application of these frameworks and arrangements will work to minimise most of the potential risks posed by the 10 PIA Topics covered by this PIA Update, including the most significant potential privacy risk, being the risk of reidentification of data subjects. This is an existing and inherent risk that is continuously managed as part of existing PLIDA processes, including the use of privacy threshold assessments (**PTAs**) and separate privacy impact assessments (**PIAs**) as specific changes or issues are identified, which we support as being consistent with privacy best practice.
- 5.2 A table summarising compliance risk of the PIA Topics against each APP is set out in **Part D [Summary of Privacy Compliance]**.
- 5.3 While we have not identified any privacy compliance risks associated with the PIA Topics that do not already have existing and appropriate mitigation strategies in place, we have identified the following issues that we consider warrant further attention:
- 5.3.1 if personal information will be received from NFP organisations for inclusion in PLIDA, there is a risk that a particular NFP organisation may not have the capability, interests and/or resources needed to properly exercise the important functions of a Data Custodian under the existing PLIDA processes and governance framework – which in turn may impact on the effectiveness of some of the mitigation strategies generally used to reduce privacy risks for PLIDA (PIA Topic 2);
 - 5.3.2 uncertainty around existing requirements for a separate PIA process to be undertaken before sensitive information or ‘inherently sensitive’ information, which was originally collected by a Data Custodian via a survey process, can be included in PLIDA (PIA Topic 3);
 - 5.3.3 the potential for *perception* of a heightened reidentification risk around some expanded outputs from PLIDA, particularly the creation of some modules, and family and household structures centred on data subjects who are children that will be able to be accessed by researchers as part of the LCDI (PIA Topic 8); and
 - 5.3.4 the need for continual (and potentially enhanced) transparency around the operation of PLIDA, its datasets, and approved research projects (PIA Topics 1, 2, 3, 4, 5, 6, 8 and 9).
- 5.4 These risks have been further considered throughout this PIA Update report. The best practice recommendations set out in paragraph 6.1 are designed to address the identified risks and further enhance privacy protections, and/or further strengthen the ABS’ compliance with the APPs.

6. Recommendations

- 6.1 This PIA Update makes the following **best practice recommendations**² in relation to the issues covered as part of this PIA Update:

Recommendation 1 NFP organisations as Data Custodians

Rationale

Data Custodians play a pivotal role in PLIDA, including to ensure that data can properly be provided to PLIDA. Data Custodians also play an important role when they assess proposed research projects, by ensuring that reidentification risks in relation to PLIDA data used for the specific project have been appropriately considered and will be properly mitigated as part of that project. However, NFP organisations will have varying levels of resources and expertise (noting that many NFP organisations are not subject to the Privacy Act), and they may not have obligations to take into account the same public policy/public interest considerations as Data Custodians who are government entities. Therefore, it is not clear whether all NFP organisations will be best placed to meet the PLIDA Board's expectations for discharging the responsibilities of a Data Custodian in relation to any data that they supply for inclusion in PLIDA.

Recommendation

We **recommend** that the ABS build upon its existing processes that are used to consider whether new datasets should be included in PLIDA³ so that, when consideration is being given to including data in PLIDA from an NFP organisation, consideration is also given to:

- whether there is an appropriate level of benefit in obtaining the data directly from that NFP organisation (as opposed to, for example, receiving that data from another entity, such as a government agency to which the NFP organisation already provides the data outside of PLIDA), including the quality of the data that would be received;
- whether it is appropriate for that NFP organisation to be the Data Custodian for that data, and exercise the rights and responsibilities of the Data Custodian under the PLIDA governance framework. This process may require consideration of:
 - the NFP organisation's overall privacy maturity, and its readiness or willingness to comply with privacy obligations in relation to the relevant data (e.g. the NFP organisation may or may not already be an 'APP entity' under the Privacy Act; or be willing to 'opt-in' to obligations under the Privacy Act⁴; or have equivalent obligations under State or Territory privacy law); and
 - whether the NFP organisation can practically take on the responsibilities of an effective Data Custodian, such as:
 - the technical expertise of its personnel (e.g., whether the NFP organisation already employs data managers, who have ability and capacity to appropriately consider reidentification and other risks for specific research projects);
 - whether it has an appropriate data governance framework that it would apply when considering specific research projects; and
 - the NFP organisation's ICT arrangements (including its ability to safely transfer data between the NFP organisation and the ABS).

² 'Best practice' recommendations are measures that promote a stronger privacy posture within the ABS in relation to PLIDA, and are distinct from 'compliance' recommendations (which are designed to meet minimum compliance requirements under the relevant APP).

³ See: <https://www.abs.gov.au/about/data-services/data-integration/integrated-data/person-level-integrated-data-asset-plida/privacy-plida>.

⁴ For example, under section 6EA or 6F of the Privacy Act.

Recommendation 2 Approach to determining ‘inherently sensitive survey data’

Rationale

Some information that does not fall within the meaning of ‘sensitive information’ under the Privacy Act nevertheless has a degree of inherent sensitivity, in the sense that data subjects may have very serious concerns if it was collected or handled in an inappropriate manner (for example, a person’s experience of domestic or family violence). In many cases, best practice (and Australian community expectations) may suggest that data from which such information can be ascertained should be afforded additional protection, even if this is not required by the Privacy Act. However, developing a comprehensive definition for this type of information is difficult.

If data that was originally collected by a Data Custodian via a survey mechanism contains ‘sensitive information’ or ‘inherently sensitive’ data, existing PLIDA processes require a separate PIA process to be undertaken by the ABS in respect of the proposed handling of that data. However, there is currently no PLIDA guidance or processes for deciding whether such data should be considered ‘inherently sensitive’.

Recommendation

We **recommend** that the ABS retain its existing PLIDA process of ensuring that a PIA is conducted before including in PLIDA any data which was originally collected by a Data Custodian via a survey mechanism, and contains ‘sensitive information’ (as defined in the Privacy Act) or ‘inherently sensitive survey data’. It would be useful if the ABS could document and publish this process, including further information about the role of the Data Custodian in any such decision (and in any PIA process itself).

However, we also **recommend** that the ABS develop guidance to assist in determining whether survey data should be considered as ‘inherently sensitive’. Given the difficulties in formulating a comprehensive definition, we instead suggest that this guidance should document factors to be taken into consideration when making such a determination. We consider the key factors should be whether inappropriate release of the data (in connection with an identified person) would be likely to result in harm to that person (such as discrimination or embarrassment), and if so, what level of harm would be caused. This test could potentially reflect the factors that inform whether a data breach is likely to cause ‘serious harm’ under Part IIIC of the Privacy Act (such as the potential for identity theft, financial loss through fraud, or physical, psychological or reputational harm). Details of the survey itself (e.g. time since the survey and the subject matter of the survey) will also be relevant in determining whether the test has been met.

The ABS could also include in the guidance a non-exhaustive list of data items that should be considered ‘inherently sensitive’ for the purposes of PLIDA. We suggest any such list could include:

- data about exact income (noting that data about income ranges may not be considered ‘inherently sensitive’ if sufficiently broad ranges were used);
- data about citizenship and visa status in some cases (some visa subclasses may reveal information relating to ethnicity, or refugee status)
- data about any experience or consequences of family, domestic or sexual violence; and
- data about unemployment.

Recommendation 3 Justifying combination of the ‘Linker’ and ‘Librarian’ roles within the ABS

Rationale

A change to the existing roles within ABS may be perceived as a reduction of privacy protections for data within PLIDA.

Recommendation

If the ABS decides to proceed with this change, we **recommend** that it prepare and publish information that explains:

- the reasons for the change (including the historical reasons why the two roles were originally implemented);
- how the change now represents best practice (noting that the change aligns with standard practice in other accredited integrating authorities, and has been implemented by the ABS in respect of other integrated data assets); and
- how the Five Safes Framework and the separation principle will continue to be implemented in respect of data within PLIDA following the change.

This could include publication of the risk assessment that the ABS intends to conduct about the combined Linker/Librarian role, which would provide greater assurance to stakeholders about the continued application of the Five Safes Framework (and, particularly, continued application of the separation principle).

Recommendation 4 Use of ANDII ICT solution for PLIDA processing and to ‘re-use’ PLIDA data for the NDDA

Rationale

Openness and transparency in handling personal information is key for the ABS to build and maintain public trust in the continued use of PLIDA, and an important aspect of its obligations under the Privacy Act. We identified some confusion or misunderstandings among stakeholders about the differences between the PLIDA and NDDA data governance arrangements, and the implementation of the proposed changes to the ICT infrastructure that will be used for PLIDA.

Recommendation 4.1: We **recommend** that the ABS develop (and if appropriate, publish) information that:

- explains the nature and benefits of the change to use the ANDII ICT solution for PLIDA data integration work (e.g. that the change will significantly improve the efficiency of work to prepare and process large amounts of data);
- explains that all PLIDA data will continue to remain within an ABS-controlled ICT environment (until outputs have been vetted, noting that at this point data can exit the DataLab);
- clarifies that despite the ANDII ICT solution also being used to handle data for the NDDA, any PLIDA data in the ANDII ICT solution will be stored and handled separately to NDDA data and remain subject to the PLIDA governance framework;
- explains the security protections that will apply to the ANDII ICT solution, which will work to protect PLIDA data (in a manner at least equivalent to the ICT systems currently used by the ABS); and

- clarifies, particularly for Data Custodians, that this change will not involve any change to the mechanisms currently used to authorise the collection and use of PLIDA data (to avoid an assumption that the change involves a move to reliance on the *Data Availability and Transparency Act 2022* (Cth) (**DAT Act**)).

Recommendation 4.2: We **recommend** that the ABS develop (and if appropriate, publish) information that explains how PLIDA data can be ‘re-used’ for the NDDA, including that:

- the Data Custodian under the PLIDA arrangement must approve the use of the PLIDA data for the NDDA, and this approval must be documented in an appropriate data sharing arrangement under the NDDA governance framework;
- there will effectively be two copies of the same analytical data (these may be identical or one may be a subset of the PLIDA data, and/or one copy may have been treated with additional data suppression to meet NDDA requirements) – and that one will be stored and used in relation to PLIDA, and the other separately stored and used for the NDDA; and
- the NDDA governance arrangements will only apply to the NDDA copy of the data (all PLIDA arrangements, including for the use of data for approved research projects, will continue for the PLIDA copy of the data).

Recommendation 5 Governance arrangements for the Core Modules and the LCDI

Rationale

As part of our consideration of the creation of the new Core Modules and the LCDI, we identified a potential *perception* that the creation and use of the Core Relationships Module and/or the family and household structures for the LCDI may involve an enhanced reidentification risk, and this perception may be amplified because these data tools facilitate examination of changed relationships of data subjects (particularly children) over time.

While we are satisfied that the ABS has taken appropriate steps to mitigate reidentification risks (including through careful consideration of the Core Modules through Safe Data Risk Assessments and review of particular aspects by ABS Disclosure Review Committee, and the application of the existing strategies in the PLIDA governance framework for access by Authorised Users of the Core Modules), the lack of publicly available information about the Core Modules may mean that a potential *perception risk* remains.

This potential *perception risk* also arises in relation to the creation of the family and household structures for the LCDI. Again, while we consider that the existing PLIDA governance framework in place should be sufficient to mitigate any reidentification and other privacy risks for the inclusion of new datasets into PLIDA for use in connection with the LCDI, and the creation and use of the family and household structures in connection with the LCDI, we found a lack of documentation that properly explains the LCDI and how privacy risks have been considered and addressed.

Recommendation 5.1 (new Core Modules): We **recommend** that the ABS develop, and if appropriate publish, information that explains:

- that ‘Safe Data Risk Assessments’, including, where relevant, consideration by the Disclosure Review Committee, have considered the reidentification risks associated with the creation and use of the Core Modules, including the Core Relationships Module, and concluded that these risks are low and manageable – this will assist in explaining that these data tools do not involve any enhanced reidentification risk when used within the existing PLIDA governance framework; and
- how Authorised Users that will use the new Core Modules will ensure that use of the relevant Module is appropriate for the particular research project (for example, if the Core Indigenous Module or the Core Relationships Module will be used for projects to examine data about a particular cohort such as Aboriginal and/or Torres Strait Islander peoples,

when a cultural review panel will be consulted, or other cultural governance and testing will be undertaken).

Recommendation 5.2 (LCDI)

We **recommend** that the ABS:

- consider undertaking 'Safe Data Risk Assessments' in respect of the LCDI (particularly the creation and use of the family and household structures), and arrange for the Disclosure Review Committee to consider the reidentification risks associated with the creation and use of the family and household structures in connection with the LCDI – this will assist the ABS to demonstrate that these risks have been carefully considered, and are determined to be low and manageable; and
- consider developing, and if appropriate publishing, information that:
 - explains how the reidentification risks have been assessed in relation to the LCDI (to explain that the LCDI and its new data tools do not involve any enhanced reidentification risk when used within the existing PLIDA governance framework);
 - explains how Authorised Users that will be able to use the family and household structures will ensure that use of the relevant Module is appropriate for the particular research project; and
 - clarifies that the LCDI will not result in a single integrated dataset being maintained by the ABS (or all data within such a dataset being accessible by all LCDI researchers); but rather that the LCDI involves new datasets being added to PLIDA, with specific research projects which would result in access to these and other PLIDA datasets being considered in accordance with the existing PLIDA governance framework (which will carefully consider all relevant factors, including the public benefit of the research and any potential reidentification risks).

Recommendation 6 Additional measures for approved research projects to support surveys

Rationale

While we have concluded that existing mitigation strategies are likely to be sufficient to address privacy risks associated with approved PLIDA research projects that will support surveys undertaken by government entities, privacy will be enhanced if the data subjects who participate in the relevant survey are made aware (and therefore should reasonably expect) the use of other data about them in connection with the relevant survey.

Recommendation

We **recommend** that the ABS continue to:

- work with Data Custodians (including relevant parts of the ABS that conduct surveys, including for the Census and ABS Household Surveys) to ensure that their survey processes are designed so as to increase the likelihood of data subjects understanding how their survey answers may be supplemented by, or used in connection with, other data that they have provided to other entities (particularly to government agencies). This may include reviewing and/or updating APP 5 collection notices and consent processes to reflect the proposed incorporation into, and use of, PLIDA data in connection with that survey; and
- publish information about approved PLIDA projects, including examples of the approved research projects that support surveys, as an openness and transparency measure.

Part B Methodology and Assumptions

7. Our methodology

- 7.1 This PIA Update has been conducted to identify and consider any new or enhanced privacy impacts and risks associated with the PLIDA changes and new functionality within the scope of this PIA Update.

Stage	Description of steps
1.	<p>Plan for the PIA Update: We reviewed relevant background material provided by the ABS (listed in Attachment 1 [Material Reviewed]), and were provided with a briefing by ABS staff.</p> <p>We discussed the various changes that have been made, or are being considered to be made. To assist in ensuring that we had correctly understood the background and to increase awareness of the issues we considered were likely to be important for this PIA Update, we prepared an Issues Guidance document to assist in the design of the relevant stakeholder engagement process.</p>
2.	<p>Project Description: We prepared an initial draft Project Description, which described our understanding of each of the topics within the scope of this PIA Update (as set out in paragraph 2.1 in Part A [Executive Summary]). This draft was refined and then finalised following feedback from the ABS.</p>
3.	<p>Stakeholder Consultation: Undertaking consultation with stakeholders was considered an essential part of conducting this PIA Update. It provided an opportunity to inform stakeholders about the topics within the scope of this PIA, and to listen to stakeholder views about them.</p> <p>We assisted the ABS to prepare and plan stakeholder consultation workshops, including advising on stakeholder consultation lists and commenting on draft material for the consultation. Our draft Project Description from the previous step was provided to stakeholders prior to the consultation sessions as background material to inform the discussions.</p> <p>Stakeholder consultation sessions for this PIA Update involved participation by 32 organisations. We attended and participated in the consultation sessions, before preparing a report summarising the feedback provided by stakeholders during the consultation sessions. We produced a Consultation Report, which summarises the feedback received. A copy of the Consultation Report will also be published on the ABS' website.</p>
4.	<p>Privacy impact analysis and compliance check:</p> <p>In this step we focussed on compliance against each relevant APP and privacy best practice. In undertaking our analysis we considered and applied <i>Australian Privacy Principles Guidelines (APP Guidelines)</i> issued by the OAIC, which outline the mandatory requirements of the APPs, how the OAIC will interpret the APPs, and matters that may be taken into account when assessing the ABS's compliance with the Privacy Act.</p> <p>In addition to the valuable insights gained through the stakeholder consultation process, we also used our knowledge of Australian community expectations gained from research and related work in identifying and analysing privacy risks.</p> <p>Where we identified a privacy risk for a particular PIA Topic, we allocated a rating that reflects our assessment of the compliance risk against the relevant APP(s), based on the following descriptors:</p>

Stage	Description of steps	
	Compliance Rating	Description of Compliance Rating
	Significant Risk	This rating indicates that the recommendation is made to address a high compliance or other privacy risk, where we consider the measure must be undertaken before the project proceeds further, to ensure compliance with the relevant APP(s) and best privacy practice.
	Compliance Risk	This rating indicates that the recommendation is made to address a moderate compliance or privacy risk, which we consider should be addressed in order to further enhance compliance and/or the privacy protections for individuals.
	Improvements to meet best practice	This rating indicates the recommendation is made to address a low privacy risk, but where the measures are still recommended to meet privacy best practice.
	Compliant	This rating indicates that we have concluded that no further mitigation steps are required.
	A summary of overall application of compliance risk ratings against the APPs for all PIA Topics is set out in Part D [Summary of Privacy Compliance] of this PIA Update report.	
5.	Privacy management and addressing risks: We considered potential mitigation strategies that could reduce or remove the privacy impacts and risks identified during the previous step, and developed our recommendations.	
6.	Draft report: We prepared a draft version of this PIA Update report.	
7.	Further refinement of draft PIA Update report: Following review of the draft report by the ABS and the PLIDA Board, we further refined our analysis and potential mitigation strategies as required to ensure that privacy risks were appropriately considered and addressed.	
8.	Final report: We finalised this PIA Update report.	

7.2 We understand that the PLIDA Board will review this PIA Update report and separately respond to our recommendations.

8. Assumptions and qualifications

- 8.1 The scope of this PIA Update report is limited to the issues and potential changes to PLIDA as set out the description and information flows in **Part C [Overview of PLIDA]**, **Part E [Privacy Analysis – Inclusion of new and expanded data types in PLIDA]**, **Part F [Implementation of new or changed data handling practices for PLIDA]** and **Part G [Privacy Analysis – Expanded outputs from PLIDA]**.
- 8.2 We have conducted our analysis on the basis that the factual information provided by the ABS (as set out in **Part C [Overview of PLIDA]**, **Part E [Privacy Analysis – Inclusion of new and expanded data types in PLIDA]**, **Part F [Implementation of new or changed data handling practices for PLIDA]** and **Part G [Privacy Analysis – Expanded outputs from PLIDA]**) is up-to-date, correct and complete.
- 8.3 Some of the topics that are within the scope of this PIA have already been considered by previous and/or separate PTA or PIA processes. We have not in this PIA Update report endeavoured to replicate the consideration given in those previous processes, but rather have used those PTA/PIA reports to identify any ‘gaps’ that represent a previously unidentified privacy risk. We have not confirmed that recommendations arising from any previous PTA/PIA processes have been implemented by the ABS.
- 8.4 As at the date of our analysis for this PIA, some of the proposed reforms to the Privacy Act (including those in the *Privacy and Other Legislation Amendment Act 2024* (Cth)) had not yet commenced, or had not yet been prepared and introduced into Parliament⁵. While we have endeavoured to take all proposed reforms into account, our analysis is based upon the provisions of the Privacy Act, and associated case law and guidance material as at the date of analysis on the cover page of this PIA Update report. However, we support the existing PLIDA processes that will ensure that privacy impacts for PLIDA continue to be considered over time (including through privacy threshold assessments, privacy reviews undertaken by the ABS, and additional or supplementary PIA processes), which will work to address risks arising about of proposed reforms to the Privacy Act and the inherent privacy risk of ‘function creep’ (which is an inherent privacy risk for projects that involve personal information being held and used over a significant period of time).

⁵ Other proposed reforms include those discussed the *Privacy Act Review Report* released by the Attorney General's Department (available at <https://www.ag.gov.au/rights-and-protections/publications/privacy-act-review-report>), and the Australian Government's response (available at https://www.ag.gov.au/rights-and-protections/publications/government-response-privacy-act-review-report#:~:text=In%20its%20response%20to%20the,to%20best%20protect%20this%20information.)).

Part C Overview of PLIDA

9. PLIDA Governance

- 9.1 PLIDA was established through a cross-portfolio government partnership of Commonwealth agencies, including the ABS, the Australian Taxation Office, Department of Education, Department of Health and Aged Care, Department of Social Services, Services Australia and Department of Home Affairs.
- 9.2 The ABS is the accredited integrating authority for PLIDA and is responsible for:
 - 9.2.1 combining the datasets in PLIDA;
 - 9.2.2 providing access to PLIDA data to only those individuals (**Authorised Users**) from entities who have been authorised to access the PLIDA data for approved research projects;
 - 9.2.3 ensuring the security of data contained within PLIDA; and
 - 9.2.4 ensuring that all research results and other outputs from the use of PLIDA data are produced in a manner that is unlikely to enable the identification of a particular person.
- 9.3 PLIDA is bound by the constraints of:
 - 9.3.1 the legislation of the Data Custodians that applies to data that they provide to PLIDA;
 - 9.3.2 the *Census and Statistics Act 1905* (Cth) (**Census and Statistics Act**);
 - 9.3.3 the Privacy Act; and
 - 9.3.4 social licence from, and the community expectations of, Australians.

10. Benefits of PLIDA

- 10.1 The analytical data in PLIDA is used by Authorised Users for approved research projects in a secure, virtual access environment. PLIDA presents benefits to Australian agencies, the Australian public, researchers, and Data Custodians by:
 - 10.1.1 providing Australian governments with a powerful tool for informing government decision-making;
 - 10.1.2 supporting decisions that will help Australians live healthier, happier, and more independent lives;
 - 10.1.3 making better use of the information that has already been collected to enhance the value of existing public data resources; and
 - 10.1.4 making a wider range of data available for researchers from government, universities, and public policy institutes.

11. PLIDA data

- 11.1 PLIDA contains high-value, person-centred and regularly updated datasets that aim to comprehensively cover the Australian population. The data contained within PLIDA is currently primarily made up of Commonwealth data that is provided to the ABS by Data Custodians (or other entities authorised by Data Custodians to provide the data). Where the relevant individuals in those datasets are able to be identified, the datasets will include personal information and sensitive information (as defined in the Privacy Act). The ABS has compiled a description of the datasets that are currently included in PLIDA, and this is set out in **Attachment 4**.
- 11.2 The ABS has built a central linkage infrastructure called the 'Person Linkage Spine' (**Spine**) which represents the Australian population. PLIDA is built by linking the datasets supplied by Data Custodians to the Spine, instead of linking each dataset to every other dataset individually. This creates the capacity for separate source datasets to be linked to one another via the Spine, in a flexible and efficient manner. Through the Spine, the datasets can be combined as required and extracts of the linked data can be used by Authorised Users, such as researchers, for approved research projects. The Spine enables information to be brought together in relation to people who were resident in Australia during a given reference period from 2006 to present. Diagrams illustrating the overarching data flows for PLIDA in its current state are included in **Attachment 2**, with its proposed state in **Attachment 3**.
- 11.3 The ABS is responsible for ensuring that PLIDA data that it makes available to Authorised Users is provided in a manner that is not likely to enable the identification of an individual (and therefore meets the requirements to be 'deidentified' under the Privacy Act).⁶
- 11.4 The ABS currently has a range of security arrangements in place for the IT systems that are used for PLIDA to protect PLIDA data, which:
- 11.4.1 conform with security arrangements set out in the Australian Government Information Security Manual (**ISM**);
 - 11.4.2 ensure that data collection, linkage, and assembly activities for PLIDA datasets are only conducted by a dedicated team in the Secure Data Integration Environment (**SDIE**);
 - 11.4.3 ensure that access to data by Authorised Users outside the ABS is conducted only by giving access to analytical data within an ABS environment (known as the **DataLab**)⁷;
 - 11.4.4 includes a secured internet gateway which is reviewed annually by the Australian Signals Directorate (**ASD**); and
 - 11.4.5 includes an ongoing program of security audits and system accreditations, including the Information Security Registered Assessors Program (**IRAP**).

⁶ Personal information is deidentified 'if the information is no longer about an identifiable individual or an individual who is reasonably identifiable' (section 6(1) of the Privacy Act).

⁷ A limited number of PLIDA activities can occur on other ABS-controlled infrastructure known as SEADpods (see 2019 PIA Update).

- 11.5 Further, PLIDA datasets are handled in accordance with a range of additional privacy protection practices, including:
- 11.5.1 the Five Safes Framework, which is an internationally recognised approach to managing disclosure risks, which is applied to ensure access to PLIDA data is appropriate. The framework is designed to facilitate safe data release using five elements (Safe People, Safe Projects, Safe Settings, Safe Data and Safe Outputs) which are all assessed independently, but also considered as a whole for each instance of data access;
 - 11.5.2 data may be added to PLIDA via a once-off linkage (for a specific research project or projects and will not be retained following the completion of the project(s)), or as a part of an enduring analytical asset that is separate to PLIDA; and
 - 11.5.3 the ABS is transparent about linkages with PLIDA and approved research projects that make use of PLIDA data, through information included on the ABS website.

Part D Summary of Privacy Compliance

12. Summary of Compliance

12.1 **Table 1** below summarises our analysis of the various PIA Topics against each APP, and allocates an 'APP Compliance Risk Rating' for each APP in accordance with the descriptors set out in the table under paragraph 7.1 in **Part B [Methodology and Assumptions]** in of this PIA report. This analysis does not address those elements of the APPs which reflect the ABS' broader compliance obligations.

Table 1: Summary of APP compliance

APP	Risk Rating	Summary
APP 1 Open and transparent management of personal information	Improvement measures to meet best practice	<p>Further transparency measures to meet best practice are recommended in relation to:</p> <ul style="list-style-type: none"> inherently sensitive survey data (see Recommendation 2); changes to the ABS roles in preparing PLIDA data (see Recommendation 3); proposed use of the ANDII ICT solution for PLIDA, and for re-use of PLIDA data for the NDDA (see Recommendation 4); the new Core Modules, and the family and household structures that will be created, and may be used, in connection with the LCDI (see Recommendation 5); publication of PLIDA projects to support surveys (see Recommendation 6). <p>Further governance measures to assist with enhancing compliance or best practice in relation to all APPs are also recommended (see Recommendation 2).</p>
APP 2 Anonymity and pseudonymity	Not relevant to PIA Topics	<p>Data subjects will not be specifically dealing with the ABS for inclusion or use of data in PLIDA.</p> <p>Other individuals (personnel within Data Custodians or Approved Users) will not be impacted by the topics covered by this PIA Update.</p>
APP 3 Collection of solicited personal information	Improvement measures to meet best practice	<p>Further measures to support the collection of data into PLIDA are recommended in relation to:</p> <ul style="list-style-type: none"> data to be collected from NFP organisations (see Recommendation 1).
APP 4 Dealing with unsolicited personal information	Not relevant to PIA Topics	<p>The requirements under APP 4 for collection of unsolicited personal information are not relevant to the topics covered by this PIA Update.</p>
APP 5 Notification of the collection of personal information	Improvement measures to meet best practice	<p>Best practice measures to promote the principle behind APP 5 are recommended in relation to:</p> <ul style="list-style-type: none"> approved research projects to support surveys (see Recommendation 6).

APP	Risk Rating	Summary
APP 6 Use or disclosure of personal information	Improvement measures to meet best practice	Further measures are recommended to address potential risks in relation to the use and/or disclosure of personal information (including the risk of disclosure as a result of reidentification) in relation to data to be collected from NFP organisations (see Recommendation 1) and the family and household structures to be used in connection with the LCDI (see Recommendation 5.2).
APP 7 Direct marketing	Not relevant to PIA Topics	The requirements under APP 7 for use of personal information for direct marketing are not relevant to the topics covered by this PIA Update (no data will be used by organisations for direct marketing purposes).
APP 8 Cross-border disclosure	Compliant	We consider that the existing mitigation strategies, and implementation of recommendations from previous PIA processes, will be appropriate to mitigate risks associated with any access to PLIDA data from outside Australia associated with the topics within scope of this PIA Update.
APP 9 Government related identifiers	Not relevant to PIA Topics	The requirements under APP 9 for adoption or use/disclosure of government related identifiers by organisations are not relevant to the topics covered by this PIA Update (there will be no use of government identifiers by organisations).
APP 10 Quality of personal information	Compliant	We have not identified privacy impacts associated with the quality and accuracy of data included in PLIDA, which will not be mitigated by existing strategies.
APP 11 Security of personal information	Compliant	We have not identified privacy impacts associated with security or retention of data included in PLIDA, which will not be mitigated by existing strategies.
APP 12 Access to personal information	Not relevant to PIA Topics	None of the topics within scope of this PIA Update change the existing arrangements in relation to access to personal information (noting that once data is in PLIDA, data subjects will not be able to access their personal information).
APP 13 Correction of personal information	Not relevant to PIA Topics	None of the topics within scope of this PIA Update change the existing arrangements in relation to correction of personal information (noting that once data is in PLIDA, data subjects will not be able to correct or change any of their personal information).

Part E Privacy Analysis - Inclusion of new and expanded data types in PLIDA

13. Introduction

- 13.1 In this **Part E [Privacy Analysis – Inclusion of new and expanded data types in PLIDA]**, we consider the privacy risks and impacts associated with PIA Topics involving the proposed inclusion of new or expanded data in PLIDA (in terms of type and volume), where that data will be prepared for linkage in accordance with existing PLIDA processes.
- 13.2 For each PIA Topic in this Part, we set out:
- 13.2.1 a description of the change to PLIDA, including the relevant information flows; and
 - 13.2.2 our privacy analysis of each change, including identification of the most relevant APPs engaged by the change, key privacy risks and potential impacts and existing mitigation strategies.
- 13.3 Although we considered the potential application of each APP for each PIA Topic, we have only included discussion about APPs that we considered particularly relevant or important for that PIA Topic.

14. PIA Topic 1: Linking high level administrative crime and justice data to PLIDA

Description and information flows

- 14.1 PLIDA includes many datasets containing information that was originally collected by a Commonwealth or State or Territory government agency as part of their functions and activities, before the agency as Data Custodian provided that data for inclusion in PLIDA (this is often referred to as **administrative data**). It includes data used for registrations, transactions and record keeping, usually during the delivery of a service. For example, information about the number of people who use Medicare from the Department of Health and Ageing is administrative data.
- 14.2 Some analytical datasets for administrative data in PLIDA already contain data from which it is possible to infer that a particular data subject has had an interaction with the justice system in Australia. For example, the data may indicate that a data subject received a social security payment or service that is only available to those leaving or entering prison.
- 14.3 There is a proposal to make some variables in administrative data which was received from Australian Government agencies as Data Custodians, available for use by Authorised Users in approved PLIDA projects. This is intended to facilitate approved research projects that are designed to enhance informed government policy making, and more efficient and targeted service delivery for those interacting with the justice system (for example, evaluation of employment support services for people who have left prison and are transitioning to work).
- 14.4 This new data will use variables such as 'Yes/No' flags (this is referred to as being only **'high level data'**). The existence of such variables can indicate:
- 14.4.1 that a data subject is or was in prison (e.g. the variable can be determined because the data shows their ABSTUDY custody allowance, custodial apprentice incentive payment, or Centrelink or another type of payment or service, was suspended or cancelled because they are or were in prison; or that they are or were no longer participating in a student or apprenticeship program because they are or were in prison);

- 14.4.2 that a data subject's partner, or someone else with a relationship to the data subject, is or was in prison (e.g. the variable can be determined from the type of payment or service that the data subject received); and
- 14.4.3 that the data subject, or someone other than the data subject, has had another type of interaction with crime or a justice system (e.g. the variable can be determined because the data subject received crisis or other payments or services related to family or domestic violence).
- 14.5 However, the data (including the new variables) will not contain more detailed information about any particular interaction with the criminal justice system. For example, it will not show the reasons why a data subject was in prison, what crime was committed (or alleged), or any other details about their interaction with the police, courts, or family and domestic violence protection or support systems.
- 14.6 Existing PLIDA processes mean that the ABS will ensure:
 - 14.6.1 relevant Data Custodians will confirm that there is legislative or other authority for the inclusion of the new variables in PLIDA, and their use for approved research projects;
 - 14.6.2 that the new variables will only be included in datasets that do not include any direct identifiers about the data subject (that is, they will be included in analytical datasets only, and not in the Spine used to link different datasets); and
 - 14.6.3 the new variables will only be able to be used for projects that have been approved under the data governance framework for PLIDA;
 - 14.6.4 outputs from those projects will be subject to existing PLIDA processes, designed to ensure that no individual can be identified from those outputs; and
 - 14.6.5 all of the usual security protections for PLIDA data will apply to those variables (including technical IT protections and data governance measures, including to restrict and monitor access).
- 14.7 Any proposal to include more detailed information in PLIDA about a data subject's interaction with criminal or justice systems (beyond that contained in the new high-level variables), or any proposal to include a dataset in PLIDA that contains (or is compiled from) police records, court or corrective services records, or any other data about a data subject's criminal history, will continue to require the undertaking of a separate PIA process. This means that any proposal to include State and Territory police, courts and prison data within PLIDA would be the subject of a separate PIA process. For example, the ABS is currently undertaking a PIA process for compilation of a new Criminal Justice Data Asset which will link certain State and Territory police, criminal courts and prison administrative data.⁸

⁸ As the privacy impacts will be considered as part of project-specific PIA processes, the 2024-25 PLIDA Update PIA will not examine any of these projects in any detail.

Discussion and analysis of relevant key privacy issues

Reidentification risks (All APPs)

- 14.8 The existing PLIDA arrangements mean that any high level data that is made available to Authorised Users will only contain analytical data (i.e. it will not contain any direct identifiers for the relevant data subjects). However, even if only binary variables (e.g. 'Yes/No' flags) are used, the high level data will indicate a data subject's various interactions with crime or a justice system, and may also indicate that they received (or did not receive) particular government payments or services.
- 14.9 If a data subject was able to be identified from high level data (either on its own or after such data is integrated with other analytical PLIDA data for an approved research project), the data will become 'personal information' and subject to the Privacy Act. In such a case, there may be serious privacy impacts for the identified individual. For example, if it became known that they have been in prison, that person might face real discrimination (including in the course of obtaining or maintaining a job), or serious reputational damage, both of which might lead to financial or psychological (and potentially associated physical) harms. Similarly, victim survivors of family, domestic or sexual violence may not wish this to become known – the risks associated with identification of victims and alleged perpetrators has been well documented.
- 14.10 In addition, if the identified person's criminal record can be ascertained from the data, this personal information would also be subject to the additional protections for 'sensitive information' under the Privacy Act.
- 14.11 This means that a key privacy issue associated with PIA Topic 1 is whether the proposed change raises any enhanced risk of reidentification of data subjects.
- 14.12 All handling of PLIDA data involves some reidentification risk. However, existing PLIDA processes already employ a range of technical and non-technical measures to ensure that this risk remains very low. Important measures include oversight by the ABS Disclosure Review Committee, assessments for inclusion and use of data being made against the Five Safes Framework, and taking steps to minimise disclosure risk to (or by) Authorised Users (including imposing restrictions on Authorised Users, and the ABS vetting all outputs from the DataLab). Notably, the ABS is currently undertaking an internal review process to assess the existing strategies used to address reidentification risks for PLIDA.⁹
- 14.13 We do not consider that the reidentification risks will be heightened simply by high level datasets being created from existing administrative data contained in PLIDA, and then being retained and made available for use in approved research projects.
- 14.14 However, it is possible that the use of high level data for particular approved research projects might involve an increased reidentification risk - that is, the risk arises from the combination of the high level data with other PLIDA data, rather than simply through the existence of the high level data itself. For example, reidentification risk is potentially higher if an approved research project is examining smaller remote geographies where it is more likely for individuals (and their unique attributes) to be generally known within the community.

⁹ See the ABS responses to the *Expanded Health Data Linkage to the Person Level Integrated Data Asset PIA Report* (May 2024), available here: <https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-impact-assessments>.

- 14.15 Despite this, we consider that the existing PLIDA measures should be sufficient to address any heightened reidentified risk. The role of Data Custodians in considering reidentification risk when approving proposed research projects (during which they can take into account the specific context and the intended use of any high level data, and impose any additional mitigation measures) will be particularly important in addressing this risk. As part of the project approval process, Data Custodians will be able to assess the vulnerabilities of the relevant cohort of data subjects, and the potential harms that could be caused if reidentification was to occur.
- 14.16 We note that some stakeholders raised the particular sensitivity of crime and justice issues within the Aboriginal and Torres Strait Islander community and the potential for data being used to further deficit narratives, especially if interpreted without an Aboriginal and/or Torres Strait Islander perspective. We consider this to be a significant issue, and we support the ABS' continued commitment to developing an data ethics framework (which is intrinsically linked to privacy issues, and questions about whether personal information *should* be used in a particular way, even if it *can*), and engagement with Aboriginal and Torres Strait Islander people to develop governance arrangements on the use of Aboriginal and Torres Strait Islander data.¹⁰
- 14.17 Although we have concluded that the existing mitigation strategies should be sufficient to ensure that high level data is not 'personal information' (so that the APPs do not apply), we have considered the principles behind the following APPs, as a privacy best practice measure.

Openness and transparency (APP 1)

- 14.18 It will be important to ensure that the ABS continues to be open and transparent about the creation and use of high level data in PLIDA, in a manner consistent with the principles underpinning APP 1. Stakeholders also expressed the view that data subjects should be made aware of how their data might be used (particularly if their consent will not be relied on to authorise secondary use or disclosure of their data, so that individuals will not have control over how their data will be used).
- 14.19 We support the ABS continuing to implement existing PLIDA arrangements to increase awareness about how data is handled for PLIDA, which we understand will also continue to be implemented in relation to high level data. These mitigation strategies include:
- 14.19.1 publishing information on its website (including information about approved research projects); and
 - 14.19.2 continuing to working with Data Custodians to improve information about PLIDA that is provided to data subjects at the point of collection of their personal information.¹¹

¹⁰ See recommendation 4 of the 2022 PIA Update, which included recommendations about governance of Aboriginal and Torres Strait Islander data.

¹¹ See recommendation 1 of the 2022 PIA Update, which included recommendations regarding the transparency around the inclusion of new types or categories of data within PLIDA, including providing notice to the public about them.

Inclusion of new analytical datasets in PLIDA (APPs 3, 6 and 11)

- 14.20 It is not intended that PIA Topic 1 will involve any new collection of data, rather it will involve consolidation of information that is already available in administrative data within PLIDA, albeit in a potentially more convenient form.
- 14.21 Despite this, when new high level crime and justice datasets are created there will be an expansion in the number of datasets (and the volume of data) stored by the ABS as part of PLIDA. The 2022 PIA Update considered in detail the privacy risks associated with the ABS maintaining an ever expanding data asset, but we consider that implementation of the recommendations from that previous PIA process will be sufficient to address any such risks in relation to the inclusion of the high level crime and justice datasets in PLIDA.
- 14.22 In this regard, we note that most stakeholders, particularly researchers, spoke positively about the potential public benefits arising from researchers being more easily able to access data of this nature for policy development, particularly given that the relevant cohort of data subjects is heavily impacted by government policy.
- 14.23 It is also privacy enhancing that the high level data will only indicate that a data subject has had an interaction with crime or a justice system, but will not include any more detailed information about those interactions - for example, it will not include information about any actual or alleged offence, or any other reasons why a data subject was in prison. Minimising the amount of data that is made available is a privacy-enhancing feature, and consistent with the principles behind APP 3. We support the ABS' plans to continue to conduct separate PIA processes before any specific information about data subject's crimes or interactions with a justice system will be included in PLIDA. This will assist in ensuring that APP 3 (and other APPs) are appropriately considered before that more detailed information is included in PLIDA.
- 14.24 Once included in PLIDA, the high level datasets will be subject to all of the existing governance processes and security protections for PLIDA.¹² We consider that these should be sufficient to ensure compliance with the principles behind APPs 6 and 11 in respect of that high level data.

Quality of the data (APP 10)

- 14.25 While minimising the amount of information in high level data is a privacy-enhancing feature (as discussed above), it also means that there may be a risk that inappropriate conclusions may be drawn by Authorised Users from that high level data.
- 14.26 Again, the role of the Data Custodian in approving proposed research projects is an important mitigation strategy in addressing this risk, since this allows consideration and assessment of:
- 14.26.1 the high level dataset's usefulness for the particular approved research project (including any quality limitations of the high level data); and
 - 14.26.2 the expertise of the Authorised Users in appropriately interpreting the high level data.
- 14.27 We observe that the ABS or Data Custodians for high level data could, if they wished, develop information about the quality of high level datasets and make it available to potential researchers, which may assist in ensuring that only suitable research projects using high level data will be proposed.

¹² These are detailed in the 2018 PIA, the 2019 PIA Update, and the 2022 PIA Update.

15. PIA Topic 2: Linking ‘not-for-profit’ (NFP) organisational data to PLIDA

Description and Information Flows

- 15.1 PLIDA already includes some data that is received from organisations who operate in a ‘not for profit’ manner, that is, they do not operate for the profit, personal gain or other benefit of particular people (**NFP organisations**).¹³ Many of these NFP organisations are funded by government(s) to conduct their operations and to collect particular data.
- 15.2 Some data originally collected by NFP organisations is subsequently provided to a government agency, and then provided to PLIDA by that Data Custodian as administrative data. Other data is already provided to PLIDA directly by an NFP organisation as the Data Custodian. For example, in 2023-24 the ABS undertook the Expanded Health Data Linkage PIA, which considered the inclusion of cancer registry data provided by State and Territory NFP organisations operating those registers, for projects to help understand patient journeys through the health care systems.
- 15.3 This PIA Update builds upon this consideration to assess the inclusion of other types of data from NFP organisations beyond health data.
- 15.4 It is only intended that data would be collected from large NFP organisations in relation to services that are funded by government, where there would be significant benefit and utility to include the data in PLIDA.¹⁴ Such data might include that from:
- 1.5.2 large childcare providers, for projects to help understand pathways through childcare facilities (including to understand impacts or the extent of any disadvantage children may have); or
 - 1.5.3 organisations who deliver services to people participating in their programs funded by government (for example, programs that provide supports to people experiencing homelessness or disadvantage by providing accommodation or housing solutions), for research projects to help understand the outcomes and pathways of participants in those programs, and/or to evaluate the effectiveness of those programs).

Discussion and analysis of relevant key privacy issues

Reidentification risks (All APPs)

- 15.5 NFP organisations often provide services to the most vulnerable individuals in society, who are usually least able to object to the collection and use of their personal information. The examples of data from NFP organisations that may be collected also include information about young children, who do not have capacity to object to the handling of their information, but for whom adverse privacy impacts may be increased because any harmful impacts may affect them throughout their lifetime.

¹³ This definition of NFP is sourced from the Australian Charities and Not-For-Profits Commission, available here: <https://www.acnc.gov.au/tools/topic-guides/not-for-profit#:~:text=A%20not%2Dfor%2Dprofit%20is,%2Dfor%2Dprofits%20are%20charities.>

¹⁴ It is not intended that datasets would be collected from any NFP organisations who are a ‘small businesses operator’ as defined in the Privacy Act (generally, an entity will be a small business operator if it has an annual turnover of \$3m or less for a financial year, although there are some exceptions to this rule).

- 15.6 It will therefore be critical, as it is with all data handled by the ABS for PLIDA, that reidentification risks are appropriately addressed in relation to datasets provided by NFP organisations, particularly given the potential for increased consequences for data subjects if reidentification was to occur.
- 15.7 However, we note that all of the existing mitigation strategies in place to protect against the risk of reidentification (see the discussion in PIA Topic 1 above) will apply to any datasets received from NFP organisations. We again consider that these strategies should be sufficient to protect against the risk of reidentification for this category of data.

Collection of data directly from NFP organisations (APP 3)

- 15.8 The ABS intends to obtain NFP data for inclusion directly from NFP organisations themselves, rather than via a government agency. As discussed in the 2022 PIA Update, it will be important for there to be a clear benefit in the NFP data being included in PLIDA, in the form in which it can be provided by the particular NFP organisation.
- 15.9 However, implementation of the 2022 PIA Update recommendation, so that there are clear processes for the risks and benefits of new datasets being included in PLIDA are put in place, means that these processes would also apply to data that is proposed to be received directly from NFP organisations. This is consistent with the views of some stakeholders, who queried whether there would be benefit in receiving data directly from NFP organisations, if that same data could be received from government agencies (because the NFP organisation will have separately been required to provide the data to that government agency).
- 15.10 Those processes should also be effective to ensure that the type of data being included from a NFP organisation is appropriate, noting that health data has already been considered in a separate PIA process (the Expanded Health Data PIA). Stakeholders did not identify any particular category of data that might be collected from an NFP organisation as being one which should never be included in PLIDA.
- 15.11 However, it is possible that some NFP organisations may be 'small businesses' or otherwise not subject to the Privacy Act (although, importantly, an organisation that would otherwise be a 'small business' will not be exempt from application of the Privacy Act if it is a 'contracted service provider' for a 'Commonwealth contract'¹⁵). An NFP organisation that is not subject to the Privacy Act may have had no obligation to obtain consent from, or provide information to, data subjects at the point of collection. They may also have no restriction on the further use or disclosure of the data that they hold (including by providing it to the ABS for PLIDA, even if there would be restrictions if the data was obtained from a different entity), which means that the protections in the Privacy Act will not operate until that data is received by the ABS.
- 15.12 Such an NFP organisation may also not have relevant processes in place, with suitably qualified and experience staff, to enable it to exercise the rights and responsibilities of a Data Custodian under the PLIDA governance framework, which relies on Data Custodians having sufficient privacy maturity to make informed decisions about submission of data for inclusion in PLIDA, and its subsequent use in approved research projects.
- 15.13 Data Custodians who are government agencies are expected to have appropriate skills and experience in weighing the public benefit and utility of data use against data sensitivity, restrictions or safety considerations (including reidentification and other privacy risks).

¹⁵ These terms, defined in the Privacy Act, effectively mean that an entity that is (or was) party to a contract with a Commonwealth agency, under which the organisation is (or was) responsible for the provision of services to an agency or State/Territory authority (or a subcontractor to such an entity), will be bound by the Privacy Act.

- 15.14 In contrast, NFP organisations with lower data maturity or insufficient understanding of privacy law may be unable to appropriately assess, and make determinations on the basis of, privacy risks associated with the use of the data they supplied to PLIDA. It is also possible that NFP organisations could have competing commercial or other interests that may impact its decision-making around research project approvals (for example, NFP organisations may not support proposals for research projects that may result in outcomes that could portray the NFP organisation unfavourably, or jeopardise future funding opportunities, even if the outcomes would otherwise be socially valuable).
- 15.15 This may affect the ability of an NFP organisation to act as an effective safeguard against the inappropriate inclusion in PLIDA, or the use of PLIDA data for suitable research projects. **Recommendation 1** is designed to address this risk, by ensuring that the processes used to decide whether or not it is appropriate to include new data within PLIDA are tailored to also consider whether any NFP organisation supplying that data is appropriate to be the Data Custodian of that data within the PLIDA framework, or whether alternative arrangements should be made in respect of that data.

Quality of data (APP 10)

- 15.16 If NFP organisations are not subject to the Privacy Act (or have not contractually agreed to comply with the Privacy Act, or equivalent State and Territory privacy legislation), there could be concerns about the quality of the data collected for inclusion in PLIDA, because there would be no requirement under APP 10 to take steps that are reasonable in the relevant circumstances to collect data which is accurate, or up-to-date. Some stakeholders did indeed observe that some data held by NFP organisations may be of lesser quality or reliability than that held by other Data Custodians (requiring greater data cleansing efforts from the ABS).
- 15.17 However, we do not consider that the source of data would necessarily reduce data quality within PLIDA, on the basis that the existing processes should be sufficient to mitigate any APP 10 (data quality) risks, given that:
- 15.17.1 the quality of the relevant data is already a consideration in whether or not it will be included in PLIDA;
 - 15.17.2 all incoming data to PLIDA is uniformly treated and sanitised by the ABS; and
 - 15.17.3 relevant information is made available to potential researchers about the quality of data that can be used for approved research projects.

Data Security (APP 11.1)

- 15.18 Some stakeholders also raised concerns that NFP organisations may not have appropriate ICT systems and other processes in place to ensure that personal information will be protected during its transfer to the ABS, for inclusion in PLIDA. This potentially raises an APP 11 risk.
- 15.19 However, we again consider that existing PLIDA processes should be sufficient for this risk to be assessed, and for any necessary mitigation strategies put in place, during the decision making stage for inclusion of data in PLIDA. However, implementation of **Recommendation 1** would assist in documenting this.

16. PIA Topic 3: Inclusion of 'sensitive' and 'inherently sensitive' survey data in PLIDA

Description and information flows

- 16.1 From time to time, a Data Custodian (including the ABS in its role as a Data Custodian) may propose that data that was obtained through a survey of data subjects (**survey data**) should be included as a dataset within PLIDA.
- 16.2 Such survey data may contain:
- 16.2.1 information that falls within one of the categories set out in the definition of 'sensitive information' under the Privacy Act, such as personal information about an identified individual's racial or ethnic origin, sexual orientation, criminal record or health, or their genetic or biometric information (**sensitive survey data**); or
 - 16.2.2 information that does not meet that legislative definition, but nevertheless has an inherent degree of sensitivity (for example, information about an individual's experience with family, domestic or sexual violence) (**inherently sensitive survey data**); or
 - 16.2.3 other personal information that does not fall into either of the above categories (**non-sensitive survey data**). Inclusion of non-sensitive survey data in PLIDA was considered as part of the 2019 PIA Update.
- 16.3 Currently, if survey data includes sensitive survey data, or inherently sensitive survey data, a separate PIA process is undertaken before that survey data is included in PLIDA and prepared for integration. For example, separate PIA processes have been undertaken to link data from the National Health Survey (the PIA was published in August 2018), the Survey of Disability, Ageing and Carers (this was covered in the 2019 MADIP PIA Update), and the National Study of Mental Health and Wellbeing (a PIA was published in November 2020).

Discussion and analysis of relevant key privacy issues

Identification of specific privacy risks for particular survey data (all APPs)

- 16.4 In our view, undertaking a PIA prior to the inclusion of sensitive survey data or inherently sensitive survey data in PLIDA is consistent with both privacy best practice and community expectation (as evidenced by strong stakeholder support of the existing processes).¹⁶ This ensures sufficient consideration is given to the additional protections such data may require before being included in, and made available via, PLIDA (**Recommendation 2**). Conducting a PIA will allow consideration of the specific survey and data types involved, and whether or not there is any enhanced reidentification risk for the specific data (even if direct identifiers are removed from the data). It would also allow consideration of other issues, such as whether the data should be collected by the ABS at all (APP 3) and whether notice should be given to data subjects about that collection (APP 5).
- 16.5 **Recommendation 2** also notes that documenting and publishing the existing processes would further enhance transparency (for Data Custodians, potential Approved Users, and the public) about these processes, which would further enhance compliance with APP 1.

¹⁶ We acknowledge there may be contrary views – for example, if the Australian Parliament has not seen fit to include a category of data within the definition of 'sensitive information', then it should not be afforded additional protections beyond those in the Privacy Act.

- 16.6 However, stakeholders noted the difficulty of defining ‘inherently sensitive’ information. We agree that it would be very difficult to formulate a comprehensive and conclusive definition for ‘inherently sensitive survey data’, particularly because the definition may vary over time or when a data item is considered in combination with different survey variables.
- 16.7 Instead, we recommend that ABS document and apply an approach in considering whether data is ‘inherently sensitive survey data’, focussing on the potential impacts that may flow from such information being disclosed (such as psychological harm and embarrassment). We propose an approach in **Recommendation 2**, which we believe would provide more clarity to the decision-making processes that will determine whether a separate PIA process will be undertaken.
- 16.8 We observe that, once included in PLIDA, the existing governance processes will apply to any inherently sensitive survey data for proposed research processes, which will mitigate the associated privacy risks, as discussed in the previous PIA processes.

Part F Privacy Analysis - Implementation of new or changed data handling practices for PLIDA

17. Introduction

- 17.1 In this **Part F**, we consider the privacy risks and impacts associated with PIA Topics involving the proposed implementation of new or changed data handling practices for PLIDA.
- 17.2 For each PIA Topic in this Part, we set out:
- 17.2.1 a description of the change to PLIDA, including the relevant information flows; and
 - 17.2.2 our privacy analysis of each change, including identification of the most relevant APPs engaged by the change, key privacy risks and potential impacts and existing mitigation strategies.
- 17.3 Although we considered the potential application of each APP for each PIA Topic, we have only included discussion about APPs that we considered particularly relevant or important for that PIA Topic.

18. PIA Topic 4: Combining the 'Librarian' and 'Linker' roles¹⁷

Description and Information Flows

- 18.1 PLIDA adheres to the 'separation principle' by implementing 'functional separation' in receiving, storing, and curating PLIDA data.¹⁸ This means for each dataset included in PLIDA, data that contains identifying information about an individual who is the subject of the dataset (linkage information), is stored and handled separately from other data about that individual (analytical data), which does not contain any direct identifiers about such individuals. Access to these different information sets is restricted so that no individual is able to access both sets of information simultaneously.
- 18.2 Currently, ABS personnel responsible for handling linkage information are categorised into two separate roles: Librarians are responsible for cleaning, standardising, anonymising and otherwise preparing data for linkage; and the Librarians then pass the data to Linkers, who undertake the necessary processes to link the data to the Spine. Under current PLIDA arrangements, the same person cannot undertake both roles. Neither Librarians nor Linkers have any access to the relevant analytical data (this is undertaken by ABS personnel performing other roles, such as Assemblers).
- 18.3 The separation of the Librarian and Linker roles is largely a historical feature, and was implemented by the ABS at a time when MADIP (as the precursor to PLIDA) was one of the first enduring national data integration projects in Australia intended to support multiple projects on an ongoing basis. Other accredited data integrators (such as the Australian Institute of Health and Welfare (AIHW) and the Australian Institute of Family Studies) do not apply a similar distinction between these two roles. Further, the ABS does not apply the distinction in Librarian/Linker roles for other projects where it is the accredited data service provider (ADSP) under the DAT Act, which sets out a robust framework based on international standards for data sharing. One such project is the NDDA, which uses a single role to perform Librarian and Linker functions.

¹⁷ In the Stakeholder Consultation Report, this topic is described as 'PIA Topic 6'.

¹⁸ Further details on the separation principle and functional separation can be found at: <https://www.abs.gov.au/about/data-services/data-integration/keeping-integrated-data-safe>.

- 18.4 Accordingly, the ABS is now considering combining Librarians and Linkers into one functional role for PLIDA, so that a single ABS personnel can perform both roles, and receive datasets containing the identifying data, and undertake hashing/unhashing processes when required. However, the ABS will still comply with the requirements for the separation principle and functional separation, by ensuring that no single personnel will access any data in PLIDA that contains both linkage information and analytical data.
- 18.5 The ABS consider that this will increase operational efficiency in performing data linkage, and allow ABS personnel to make higher quality, more informed, decisions about the linkage strategy for PLIDA.

Discussion and analysis of relevant key privacy issues

Security risks (APP 11.1)

- 18.6 APP 11 (security of personal information) is the most relevant APP in this context. In our view, combining the Linker and Librarian roles will not lead to increased privacy risk (the current level of privacy risk will be maintained). This is because, provided that the combined Linker/Librarian role does not have access to both the linkage information and analytical data, the ABS will continue its adherence to the separation principle. In practice, both the Linker and Librarian roles will only access and handle identifying information about data subjects, and never their analytical information.
- 18.7 Importantly, all ABS personnel performing the combined Librarian/Linker role will remain subject to all existing mitigation strategies designed to reduce APP 11.1 risks, including:
- 18.7.1 secrecy provisions under the Census and Statistics Act, and legal obligations of confidentiality and use of data imposed on public servants;
 - 18.7.2 restrictions on access to data stored in ABS IT systems (i.e. technical measures to implement the separation principle); and
 - 18.7.3 requirements to undertake training about their role and the handling of data for that role, which includes training on privacy obligations.

Perception of increased privacy risk (APP 1)

- 18.8 The separation principle does not require separate individuals to undertake the Librarian and Linker roles. However, we appreciate that there may be a *perception* of increased privacy risk if the distinction between Linkers and Librarians is removed, such that stakeholders may perceive that the change represents a reduction in the application of the Five Safes Framework (that is, a perception that the different roles would not have been initially implemented, if they were not an important strategy to minimise privacy risk).
- 18.9 Accordingly, we **recommend** that, if the ABS decides to proceed with this change, that it publish appropriate information which explains:
- 18.9.1 the reasons for the change (including the historical reasons why the two roles were originally implemented in the first place);
 - 18.9.2 how the change now represents best practice (noting that the unified role aligns with practices of other Commonwealth entities and by the ABS in respect of other integrated data assets); and
 - 18.9.3 how the Five Safes Framework (including the separation principle) will continue to be implemented in respect of data within PLIDA following the change

(Recommendation 3).

- 18.10 If the ABS has already conducted a risk assessment on the combined Linker/Librarian role, or proposes to do so, we observe that publication of this may also provide greater assurance to stakeholders about the continued application of the Five Safes Framework and the separation principle following the proposed change.

19. PIA Topic 5: Migrating SDIE into ANDII ICT environment for data integration purposes¹⁹

Description and Information Flows

- 19.1 In its work in establishing the NDDA, the ABS (together with the Australian Institute of Health and Welfare) has established the ANDII including the underlying ABS ICT environment (the **ANDII ICT solution**) used for processing and storing NDDA data in accordance with the NDDA governance arrangements. Currently, the ANDII ICT solution is only used for NDDA purposes, but has been designed and built for broader access-controlled use by the ABS and other agencies.
- 19.2 The ABS currently uses its Secure Data Integration Environment (**SDIE**) to undertake data integration work for PLIDA. Use of SDIE was considered as part of the 2019 PIA Update and consequentially in the 2022 PIA Update. The use of cloud-based IT services for PLIDA analytical environments (ABS DataLab and SEAD) has been considered in various other PIA processes.²⁰
- 19.3 Subject to approval of the relevant Data Custodians (and internal ABS approvals), the ABS intends to transition from use of SDIE to undertake data integration work, to use of the ANDII ICT solution for that work. For clarity, PLIDA data would be stored and used only in that part of the ANDII ICT solution that is within the ABS' IT environment (i.e. controlled solely by the ABS).
- 19.4 There will be no changes to the data integration work that is currently undertaken, or the roles within the ABS who will undertake that work – the change is to the underlying IT infrastructure and environment that will be used to undertake that work.
- 19.5 Privacy consideration of the ANDII was undertaken as part of a PIA process for the NDDA and ANDII in 2023 (the **NDDA PIA**).
- 19.6 The ABS considers that the technical functionality and security of ANDII ICT solution is at least equivalent to, and in many cases better, than the SDIE, and will significantly improve the efficiency of work that needs to be done to prepare and process the large amounts of data needed for data integration projects (i.e. building and maintaining the Spine and ensuring accurate data linkage for approved research projects).

Discussion and analysis of relevant key privacy issues

Collection, use and disclosure of PLIDA data (APPs 3 and 6)

- 19.7 In our view, the proposed change does not introduce any new privacy risks about the collection, use or disclosure of PLIDA data. While it is a technology change to the ICT environment that will be used to handle PLIDA data, it does not impact the way data is collected or handled by the ABS or Authorised Users.

¹⁹ In the Stakeholder Consultation Report, this topic is described as 'PIA Topic 7'.

²⁰ For example, the 2019 PIA Update and the 2022 Cloud DataLab PIA both considered the use of cloud services for the ABS DataLab, and the 2022 PIA Update considered further enhancements to a cloud environment to support access by approved researchers.

- 19.8 If implemented, PLIDA data used for PLIDA projects will remain logically separated from all other data stored in the ANDII ICT solution (including NDDA data) at all times, and will continue to only be handled in accordance with the PLIDA governance framework. This separation should assist in ensuring that analytical PLIDA data is not combined with non-PLIDA data (potentially increasing reidentification risks) without approval under the PLIDA governance framework.²¹
- 19.9 To the extent that this change involves any migration of personal information from the current to a new ICT system (i.e. moving identified PLIDA data, including in the Spine), we note that there is an argument that such a use would fall within the primary purpose for which the personal information was collected by the ABS (that is, to facilitate data linkage for approved research projects in accordance with the PLIDA governance framework), and therefore migration to a new ICT system to facilitate that use is consistent with APP 6. However, even if the use for migration is interpreted as being a use for a secondary purpose, we consider that such a secondary use can be authorised under APP 6.2(a). This is because:
- 19.9.1 individuals would reasonably expect that the ICT systems on which the ABS stores and handles data will be continually upgraded and replaced, which foreseeably involves migrating data between systems in order to do so; and
- 19.9.2 the use of PLIDA data for the migration can be justified as directly related to the primary purpose for collection (i.e. ensuring that PLIDA data is properly transferred to and stored on a system that allows proper access and use for the primary purpose, has a sufficiently close nexus with that primary purpose).

Security of data (APP 11.1)

- 19.10 The use of the ANDII ICT solution to appropriately protect data has been considered in other recent PIA processes (including the NDDA PIA). As the ABS is relying on the technical robustness of the ANDII ICT solution over the SDIE to justify the change, it will be important for the ABS to be able to demonstrate the extent of the security measures currently in place to protect PLIDA data (noting that some time has passed since the NDDA PIA).
- 19.11 The ANDII ICT solution has been IRAP assessed to a PROTECTED level with both the technical and operational processes in scope. The SDIE was also IRAP assessed 18 months ago to a PROTECTED level. While there has been no formal review to compare the two, the ABS has concluded that the technical and non-technical measures in the ANDII ICT solution are at least equivalent, or better, than those in SDIE, thus APP 11.1 should be satisfied.

Openness and transparency (APP 1)

- 19.12 We observe that some stakeholders consulted about this potential change seemed to incorrectly conflate the ANDII ICT solution with its use for the NDDA, which caused confusion around whether the legislative basis for collection of PLIDA data would change. Such confusion may also raise concerns about linkage activities being undertaken by other accredited data service providers for projects using the legislative authority of the DAT Act (e.g. in connection with the NDDA).

²¹ In our view, there may be greater risks of inadvertent errors being made if different governance arrangements were to be applied to a single copy of the relevant PLIDA data (e.g. there may be difficulties in establishing and implementing the necessary access and approval arrangements).

- 19.13 However, the ABS will continue to collect and use PLIDA data for PLIDA projects under the existing PLIDA governance arrangements and the Census and Statistics Act. It is not currently proposed to use other laws such as the DAT Act for PLIDA projects.²²
- 19.14 To maintain openness and transparency around data handling, it will be important for the ABS to be transparent about the proposed change, the protections in ANDII ICT solution, and the minimal privacy impacts of this technical change. Our **Recommendation 4.1** is designed to address these issues.

Quality of data (APP 10)

- 19.15 For completeness, we note that any project involving migration of large amounts of data from one ICT system to another always poses a risk that the data may be lost, corrupted or changed during transition, which in the case of personal information can involve APP 10 risks.
- 19.16 However, we do not consider that such risks are likely to be of concern for this topic, given that:
- 19.16.1 we understand that the ABS would deploy robust data migration and testing plans, to confirm the data is being moved to the right location, and ensure the quality of the PLIDA data does not change during the migration process; and
 - 19.16.2 even if PLIDA data were to be changed or become inaccessible during the migration process, there would be no consequences for any data subjects as a result.

20. PIA Topic 6: Reusing PLIDA Data for the NDDA/ANDII²³

Description and Information Flows

- 20.1 NDDA brings together deidentified data from different government agencies about Australians to help understand outcomes for people with disability. The NDDA uses authorisations under relevant legislation (including the DAT Act) to include datasets within the NDDA. Under the NDDA governance framework, agencies providing data enter into data sharing arrangements with the ABS to facilitate the handling of NDDA data, which may include a data sharing agreement under the DAT Act.²⁴
- 20.2 The ABS is accredited through the DAT Act as an accredited data service provider. The accreditation process has considered (among other things) the ABS' policies, practices, skills, and capability, as well as its data management and governance policies and practices, for performing complex data integration services, de-identification data services, and secure access data services to support data sharing.

²² See PIA Topic 6 below for discussion about re-use of PLIDA data for the NDDA, which may involve data sharing under the DAT Act.

²³ In the Stakeholder Consultation Report, this topic is described as 'PIA Topic 8'.

²⁴ Although previous PIA processes have considered the possibility of PLIDA Data Custodians using the DAT Act authorisation to provide data to PLIDA (e.g. in the 2022 PIA Update), to date consideration of the use of DAT Act authorisations in connection with the disclosure of PLIDA data has not been undertaken.

- 20.3 In 2024, various PLIDA Data Custodians entered into data sharing arrangements to give the ABS authority to provide copies of PLIDA datasets to the NDDA. This resulted in copies of the following PLIDA datasets being migrated to the ANDII ICT solution for inclusion in the NDDA:
- 20.3.1 Data Over Multiple Individual Occurrence Centrelink Administrative Data (**DOMINO**), which includes information about people who have interacted with Centrelink, including demographics, eligibility for Centrelink benefits and actual benefits received;
 - 20.3.2 Medicare Benefits Scheme (MBS);
 - 20.3.3 Pharmaceutical Benefits Scheme (PBS);
 - 20.3.4 National Disability Insurance Scheme (NDIS), including supports and outcomes;
 - 20.3.5 Survey of Disability, Ageing and Carers (SDAC); and
 - 20.3.6 Death Registrations data, which includes information about persons who have died in Australia such as their cause of death and demographic information.
- 20.4 To date, the process for inclusion of copies of PLIDA data in the NDDA has involved the secure transfer of copies of the PLIDA analytical datasets, and relevant linkage information, from the SDIE to the ANDII ICT solution (that is, the data stays within an ABS-controlled IT environment at all times). If the changes discussed in paragraphs 19.1 to 19.6 above are implemented, this process would be simplified, so that a copy of the relevant data will be created and then moved and stored within another ABS-controlled component of the ANDII ICT solution, where it will be then handled in accordance with the NDDA governance framework.
- 20.5 This process is intended to ensure that PLIDA Data Custodians who agree that copies of PLIDA data can be included within the NDDA do not have to send two copies of the same raw data to the ABS, which would involve additional security and data quality risks.

Discussion and analysis of relevant key privacy issues

'Double handling' of PLIDA data (APP 10 and 11)

- 20.6 Similarly to the proposed change discussed in PIA Topic 5, this proposed change relates to the technical handling by the ABS of data that has previously already been included in PLIDA. From a privacy perspective, we support streamlining the data transmission processes for Data Custodians who agree that the PLIDA data that they have already provided to the ABS for PLIDA may be copied and included in the NDDA. This is predominantly because requiring Data Custodians to send two (separate) copies of the same data to the ABS – one for use in PLIDA and the other for use in the NDDA – potentially raises:
- 20.6.1 security risks, in that each time raw data is re-transmitted from the Data Custodian to the ABS, each additional handling point increases the 'attack surface' of the data which may be vulnerable to unauthorised access, disclosure or misuse (and if reasonable steps could have been taken to avoid this, it may pose an APP 11.1 risk); and
 - 20.6.2 data quality risks, in that each transmission of the data from the Data Custodian to the ABS may potentially introduce inconsistencies or errors, which could require additional data cleansing efforts by the ABS (to avoid an APP 10 risk).

20.7 The proposed change means that, instead of another copy of the data being transmitted by the Data Custodian to the ABS, copies of the data will just move between ABS IT systems and therefore stay within the control of the ABS at all times, which will mitigate security and data quality risks. The NDDA copy of the data will be stored and handled within the ANDII ICT solution, separately to the existing copy of the PLIDA data (see discussion about storage in the ANDII ICT solution in PIA Topic 5 above).

20.8 We therefore consider that this proposed change will reduce APP 10 and 11 risks.

Secondary use of PLIDA data (APP 6)

20.9 Use of copies of the PLIDA data by the ABS, when it is copied and stored in the ANDII ICT solution for the NDDA, is best analysed as being a use for a secondary purpose (since the ABS originally collected it for the purposes of administering PLIDA, not the NDDA).

20.10 However, such a use would only occur if it had been authorised under the DAT Act (and the approval of the relevant Data Custodian(s) will need to have been given, in order to enter into the data sharing agreement required under the DAT Act), so that this secondary use will be permitted by APP 6.2(b).

20.11 We observe that the relevant NDDA data sharing arrangements should:

20.11.1 specify how the relevant data will be shared with the ABS for the NDDA (i.e. the 're-use' of data already held by the ABS in respect of PLIDA); and

20.11.2 specify any restrictions on the use of the data in connection with the NDDA, as required by the PLIDA Data Custodian.

Openness and transparency (APP 1)

20.12 We observe that despite our finding that there are minimal negative privacy impacts associated with this proposed change, some Data Custodian stakeholders consulted for this PIA Update voiced concerns around their potential for reduced control over PLIDA data once it was 're-used' for the NDDA.

20.13 We consider that it would be consistent with APP 1 for the ABS to develop (and if appropriate publish), information about the proposed 're-use' of PLIDA data for the NDDA to ensure there is an appropriate level of understanding about the different data assets (see **Recommendation 4.2**).

21. PIA Topic 7: International researcher access to PLIDA data in the DataLab²⁵

Description and Information Flows

21.1 The ABS is considering whether access to analytical PLIDA data (in the DataLab) can be provided to Authorised Users who are either permanently, or temporarily, located outside of Australia.

21.2 This issue was considered as part of the 2022 PIA Update, which recommended that if the ABS were to grant access to linked datasets to individuals physically located outside of Australia, that the ABS ensure that the associated risks are formerly considered and assessed as part of the project, and that additional mitigation strategies are in place.

²⁵ In the Stakeholder Consultation Report, this topic is described as 'PIA Topic 9'.

- 21.3 In accordance with that recommendation, the ABS is currently developing an approach to enable researchers outside Australia to access PLIDA data, through the conduct of trial projects. For example, it is conducting a trial in relation to an approved project involving the Organisation for Economic Cooperation and Development (OECD) and the University of Chicago, whereby researchers within these institutions are being given access to PLIDA data in the DataLab.

Discussion and analysis of relevant key privacy issues

Disclosure outside of Australia (APP 8)

- 21.4 APP 8 is relevant to the disclosure of personal information in PLIDA data to any individual located outside of Australia, even if that individual is an Australian citizen, or works for an Australian entity. It also applies if the personal information remains on servers located within Australia (as would be in the case of researchers outside Australia being given access to the DataLab). APP 8 requires that, before any such disclosure occurs, the ABS must take such steps as are reasonable in the circumstances to ensure that the overseas recipient does not breach the APPs (unless an exception in APP 8.2 applies).
- 21.5 As a starting point, APP 8 will only apply to PLIDA data that is personal information, noting that existing PLIDA processes are designed to ensure that analytical PLIDA data that is made available to Approved Users is not personal information (because the risk of an individual data subject being reasonably identifiable from that analytical data is very low).
- 21.6 However, this does not mean that the principles behind APP 8 are not relevant, particularly given that the definition of what data constitutes ‘personal information’, and the principles used to determine when an individual is ‘reasonably identifiable’, are undergoing review and change.²⁶
- 21.7 The ABS currently has a range of mitigation strategies in place to address privacy risks associated with any access to PLIDA data by international researchers (these are documented in the 2022 PIA Update). We understand that these would also be put in place in relation to any access by Australian researchers, from outside of Australia. If the relevant individual to whom access to PLIDA data will be granted has agreed to comply with the APPs in respect of the data to which they are given access, this will enhance compliance with APP 8.
- 21.8 We again note (as described in the 2022 PIA Update) that even if such an agreement is obtained by the individual located overseas, there may be some jurisdictional challenges around enforcing such a contractual obligation (or Australian privacy laws generally) for that individual (or their employer), if they improperly access or use PLIDA data, which was a concern shared by stakeholders. It is also possible that foreign researchers may misrepresent Australian data, in the absence of contextual understanding of Australian policies, demographics or regions.
- 21.9 In addition, data subjects will have no input to the governance of data once it is outside of Australia. This would particularly be the case for any data about Aboriginal and/or Torres Strait Islander peoples and communities, where data used by researchers outside of Australia may be analysed and used without being subject to Australian legislation and governance arrangements and without access to local and cultural knowledge or context. In some cases, granting international researchers access to this data may be iniquitous in circumstances where local communities struggle to gain and maintain data sovereignty. There may be some cases where researchers outside of Australia may get access to

²⁶ See the *Privacy and Other Legislation Amendment Act 2024* (Cth), and the Australian Government’s response to the Privacy Act Review Report, which noted “Importantly, the Government considers that an individual may be reasonably identifiable where they are able to be distinguished from all others, even if their identity is not known.” (page 6).

Aboriginal and/ Torres Strait Islander data to which those local communities themselves don't have access.

- 21.10 As part of the 2022 MADIP Update, the PLIDA Board previously agreed to document additional mitigation strategies for projects involving access by overseas entities and individuals. While the ABS still appears to consider DataLab access for international researchers on a case by case basis,²⁷ implementation of this recommendation (by publishing the relevant considerations and mitigation strategies) is likely to address privacy risks, and may allay any Data Custodian reluctance to provide international access to PLIDA data.
- 21.11 We also observe that the *Privacy and Other Legislation Amendment Act 2024*, which amends the Privacy Act, proposes to introduce a 'white list' mechanism under the Privacy Act, which will prescribe countries that have similar protections to the APPs. If codified into law in future, this mechanism may assist the ABS in determining which overseas recipients will be held to privacy standards requiring them to protect Australian data in a substantially similar way to the APPs. This, in turn, will mean such overseas recipients will comply with APP 8 (or the principles behind APP 8), by virtue of the exception in APP 8.2(a).

²⁷ The ABS' website advises that users must meet the ABS safe people criteria, including researchers who 'belong to an Australian organisation (international researchers and organisations will be considered on a case by case basis)'. Available here: <https://www.abs.gov.au/statistics/microdata-tablebuilder/datalab#who-can-access-the-datalab>.

Part G Privacy Analysis – Expanded outputs from PLIDA

22. Introduction

- 22.1 In this **Part G**, we consider the privacy risks and impacts associated with PIA Topics involving the proposed implementation of expanded outputs from PLIDA.
- 22.2 For each PIA Topic in this Part, we set out:
- 22.2.1 a description of the change to PLIDA, including the relevant information flows; and
 - 22.2.2 our privacy analysis of each change, including identification of the most relevant APPs engaged by the change, key privacy risks and potential impacts and existing mitigation strategies.
- 22.3 Although we considered the potential application of each APP for each PIA Topic, we have only included discussion about APPs that we considered particularly relevant or important for that PIA Topic.

23. PIA Topic 8: Creation of new PLIDA Core Modules and the Life Course Data Initiative (LCDI) ²⁸

Description and Information Flows

- 23.1 The ABS has an established process for creating PLIDA modules from PLIDA data. Modules consist of 'pre-linked' datasets which can be made accessible to Authorised Users for approved research projects. Modules are 'enduring' datasets within PLIDA (they are not created specifically for any particular approved research project), which are regularly updated and available for use in any approved research project.
- 23.2 Modules are accessible by Authorised Users in the DataLab via the PLIDA Modular Product (**PMP**), which includes the deidentified data items and reference periods for each module available to Authorised Users.
- 23.3 There is a significant demand for additional modules to be created using existing and new PLIDA datasets. These include the potential new modules described below.

New Core Modules

- 23.4 The ABS has created five new PLIDA modules (**Core Modules**) using existing PLIDA data, some of which have replaced existing modules in the PMP. The new Core Modules are:
- 23.4.1 the **Core Demographics Module**, which contains demographic items derived from data across the PLIDA datasets such as data subjects' age, gender, death dates, country of birth and highest level of education attained;

²⁸ In the Stakeholder Consultation Report, this topic is described as 'PIA Topic 4'.

- 23.4.2 the **Core Indigenous Module**, which contains information about whether or not a data subject has identified as Aboriginal and/or Torres Strait Islander.²⁹ This module replaced the 'ever identified' derived Indigenous status currently contained within the previous 'Combined Demographics – Indigenous Module'. The new module continues to use data drawn from existing administrative data within PLIDA.

This module also includes code or pseudo-code for derivation algorithms, to allow Authorised Users to construct their own Aboriginal and/or Torres Strait Islander status indicators appropriate to their research needs – it is anticipated that this module will support Authorised Users in navigating the current complexities and challenges in deriving Aboriginal and/or Torres Strait Islander status from administrative data in PLIDA.

Creation of the new module was supported by Aboriginal and/or Torres Strait Islander stakeholders separately consulted by the ABS about this change.

- 23.4.3 the **Core Locations Module**, which contains existing geographic information about data subjects within PLIDA using information from administrative data within PLIDA. It includes data items for data subjects such as their State or Territory of residence, an address register ID (**ARID**) with the start and end dates of residence at the ARID, and information about the data subject's dwelling (structure, type of dwelling, and derived residential or commercial use);

- 23.4.4 the **Core Scoping Module**, which collates information in PLIDA about data subjects' residence and activity in Australia. This module contains information about data subjects' 'vitals' (i.e. the month and year of birth and death), flags that indicate physical presence or residential status in Australia, and flags that indicate whether a data subject is in particular administrative datasets in PLIDA. It is intended that this module will better support Authorised Users to undertake projects in population scoping.

- 23.4.5 the **Core Relationships Module**, which will support research projects examining family and household analysis. It contains information about data subjects' family relationships in Australia drawn from other PLIDA datasets (such as their relationship status, start/end dates of relationship records, and relationship codes such as for 'partner' or 'parent'). This module will be used to assist with the LCDI project (discussed further below) and other projects.

- 23.5 The Spine is used to assemble the data, primarily from existing administrative data within PLIDA. However, to create the Core Modules, the ABS also needed to include some other datasets in PLIDA (using the existing PLIDA data governance framework) which were previously collected by the ABS but not previously included within PLIDA.

²⁹ We have used the phrase 'Aboriginal and/or Torres Strait Islander' to describe persons who identify as an Indigenous Australian, to respect the term that we understand most Aboriginal and/or Torres Strait Islander people prefer to use, but noting that in Australia there are many Indigenous nations, languages and cultures.

Life Course Data Initiative (LCDI) project

- 23.6 The LCDI is a pilot funded under the Targeting Entrenched Disadvantage Package, which will run from FY2023-24 to FY2026-27. One aim of the LCDI is to add new data to PLIDA, so that it can be used by particular Authorised Users (**LCDI researchers**) in selected research projects to inform long-term policy responses to disadvantage (by generating useful and actionable insights to help communities and policymakers understand and address child disadvantage). The aim of the LCDI is to improve data for research that covers the early years of Australians' lives, particularly from ages 0-14, and includes analytical information across many aspects of people's lives including health, education, employment, security, and housing.
- 23.7 The LCDI will expand the existing PLIDA datasets by adding and linking additional datasets that may include data from Commonwealth, State and Territory governments and NFP organisations (such as components of the 'ACT Health and Education', 'Better Evidence, Better Outcomes, Linked Data (BEBOLD)' and the 'Birth Registries' and 'Childcare' datasets).³⁰
- 23.8 Only selected Authorised Users who are LCDI researchers will be able to request access to the new datasets for use in approved PLIDA projects, either on their own or in conjunction with other PLIDA datasets if applicable.
- 23.9 All of the existing PLIDA processes will apply to research projects proposed as part of the LCDI, including that:
- 23.9.1 the research project must be approved by the relevant Data Custodian(s), who can impose additional restrictions, requirements or governance processes as a condition of their approval; and
 - 23.9.2 the LCDI researchers will be provided with access to the analytical data in the DataLab (and all existing requirements and restrictions on such access under the PLIDA governance framework will continue to apply);
- 23.10 To support better understanding of disadvantage as part of the LCDI, the ABS will use PLIDA (including data in the new datasets) to develop household and family structures. This can then be requested for us by LCDI researchers to facilitate their analyses in research projects on topics relating to disadvantage, particularly child disadvantage. The structures are based on data within PLIDA and created through derivations, inferences and/or imputations as described below:
- 23.10.1 **Household structures:** these will indicate all data subjects in PLIDA who live together in the same household at a particular time. The ABS will do this by:
 - (a) using the Core Locations Module to indicate the address register ID number (**ARID**) of a data subject in a given year. Authorised Users can then use the household structures to understand who cohabits (e.g. if data subjects have the same ARID at a point in time, they will be part of the same household structure at that time); and

³⁰ Some LCDI material contains references to a 'Life Course Dataset'. However, we are instructed that this is not a singular dataset that combines analytical data from various PLIDA datasets, but rather a reference to information about PLIDA datasets to which LCDI researchers can request access for an approved LCDI project, including about the new datasets which will be added to PLIDA to facilitate the LCDI.

- (b) where this is not possible (e.g. because specific address data for a data subject is not available) endeavouring to use other data in PLIDA to infer or impute the ARID for that data subject ³¹

23.10.2 **Family structures:** these will indicate, for each dependent child in PLIDA (i.e. a child aged 14 or under), data subjects who have a relationship with that child (e.g. because they are the child's parent or guardian, a partner of a child's parent or guardian³², sibling, grandparent or other family member). At times those other data subjects may be cohabiting with the child, but other times they may not be cohabiting. The ABS will do this by:

- (a) using the Core Relationships Module and other data in PLIDA that explicitly indicates a relationship between a child and other data subjects; and
- (b) where relationship information for a child is missing, endeavouring to infer or estimate relationships using regression analysis or machine learning models.

23.11 The ABS will undertake data validation activities, to check the quality of the household and family structures identified. This may involve using other datasets to see whether similar proportions of relationships are found using other sources (e.g. using released Census data to see whether similar proportions of single parent families have been identified), and examining the proportions of relationships that were created by direct, inferred or estimated methods. These activities will be used to produce documentation for LCDI researchers that describes the methods, datasets and potential quality limitations of the structures.

Discussion and analysis of relevant key privacy issues

Creation of the Core Modules (All APPs)

23.12 In our view, creation of the Core Modules, and access by Authorised Users as part of an approved research project, do not raise any privacy impacts or risks that have not already been considered as part of previous PIA processes, and which will not be appropriately mitigated by existing processes and arrangements or the implementation of previous recommendations arising out of those PIA processes.

23.13 These risks include:

- 23.13.1 privacy risks associated with the use of datasets within PLIDA that were or are needed to create the Core Modules;
- 23.13.2 potential reidentification risks associated with the creation and/or use of the Core Modules;
- 23.13.3 data quality risks associated with the use of the Core Modules; and
- 23.13.4 security risks associated with the use of the Core Modules.

23.14 We consider that these risks will be appropriately managed by the existing PLIDA processes and arrangements. In addition:

³¹ For example, other data may show that both parents of a young child have moved to a new address, so it might be inferred that it is likely that the child will also have moved to that address.

³² The data will also show if a data subject stops having that relationship with a child.

- 23.14.1 it is an important mitigation strategy that Safe Data Risk Assessments, including review by the ABS' Disclosure Review Committee where required, have been conducted to consider the reidentification risks associated with the new Core Modules, and these processes concluded that the reidentification risks were low and manageable;
- 23.14.2 we particularly support the ABS having undertaken specific consultation with Aboriginal and/or Torres Strait Islander stakeholders in relation to the creation and use of the Core Indigenous Module; and
- 23.14.3 we also recognise the importance of the ABS' plans to make sure that Authorised Users will be provided with information so they can understand the limitations of the Core Modules, and how to effectively use the Core Modules. This will be an important step in ensuring that the quality of the data in the Core Modules is taken into account, and that outcomes from different approved PLIDA research projects will be appropriate.
- 23.15 However, we were unable to locate publicly available information that explained how the potential privacy risks associated with the Core Modules have been considered, and are continually monitored as part of PLIDA governance arrangements. This may mean that a potential *perception risk* may arise. Without detailed explanation, potential concerns may arise on the expectation that the Core Modules allow approved researchers to more easily access analytical data about a specific data subject or to more easily access data which shows changes in that data subject's circumstances over time, which could be perceived as involving an enhanced reidentification risk.
- 23.16 We therefore **recommend** that the ABS consider taking steps to make more information about the Core Modules available, which better explain the steps that have been taken, and are continually taken, to address any such perception (**Recommendation 5.1**).

The LCDI

- 23.17 As PLIDA approved research projects, LCDI projects will be subject to all the existing mitigation strategies that apply to all other types of PLIDA research projects.
- 23.18 Nevertheless, we observe there may be a perception of heightened privacy risks associated with the LCDI, particularly given that some of the information we reviewed about the LCDI may give rise to some confusion or misunderstanding about the nature of the initiative (particularly references to a 'Life Course Dataset', which without further explanation could be interpreted as referring to the creation of a single integrated dataset to be used for all LCDI research projects), and noting that there was not a lot of information provided about the existing mitigation strategies that will be deployed to protect against reidentification risks.
- 23.19 Perception risks may also arise in relation to the creation of the family and household structures, which necessarily require the handling of data about data subjects who are children (given the LCDI's focus on creating child-centric data and data structures). Children do not have capacity to object to the handling of their information, and any adverse privacy impacts may be increased for children those impacts may affect them throughout their lifetime.
- 23.20 The perception risks may be heightened because the household and family structures will include information about a child data subject's relationships with other data subjects (who themselves will have unique identifiers), and their shared households. This could potentially be perceived as involving greater risk of an individual being reidentified by reference to their relationships with other people (who themselves may be reidentifiable), rather than by specific individualised attributes. If reidentification was to occur, it may result in negative impacts to both the data subjects themselves, and any identified individuals in their family or household structures (particularly in smaller geographies, where it is more likely that all members of a particular household will be known to be related within the community).

- 23.21 In addition, the perception risk may be heightened because the family and household structures will facilitate longitudinal studies over time – these studies can be perceived as a surreptitious form of ‘tracking’ if data subjects in those studies can be identified.
- 23.22 Finally, as discussed in the Stakeholder Consultation Report, stakeholders noted the difficulties in defining ‘households’ and ‘families’, including if those definitions can change over time, or if different government programs may define the concepts differently. For data subjects from an Aboriginal and/or Torres Strait Islander background, the concept of ‘family’ may not be defined in biological terms, which may introduce complexities in accurately developing family structures.
- 23.23 However, it is important to recognise that:
- 23.23.1 the family and household structures will be an extension of the Core Relationships Module – as for all of the Core Modules (as discussed above), all of the existing PLIDA governance arrangements will apply to control creation of, access to, and use of, the family and household structures (and of other PLIDA data) for LCDI research projects;
 - 23.23.2 PLIDA data already contains data about data subjects who are children, and is already used for approved research projects that are longitudinal studies - we understand that the type of data that is likely to be used in LCDI research projects has also already been included in data sharing pilots and projects conducted by other Commonwealth agencies, such as the Australian Early Development Census (AEDC) and the ATO (with Single Touch Payroll data); and
 - 23.23.3 the ABS intends that definitions for ‘family’ and ‘household’ will remain stable throughout the operation of the LCDI, while acknowledging that not all cultural nuances associated with those terms will be captured in the dataset, which may have some flow-on impacts to overall accuracy and quality of data (noting that this is an issue relevant to administrative data generally, not just that used for the LCDI).
- 23.24 With the above in mind, we consider that the existing PLIDA governance framework is sufficient to mitigate any particular sensitivities and optical risk associated with the LCDI.
- 23.25 However, it would be beneficial for the ABS to be able to *demonstrate* how the reidentification and other risks have been appropriately considered, to further aid in ensuring there is utmost transparency, and to assist in addressing any perception risks. This could be done by undertaking Safe Data Risk Assessments, including review by the Disclosure Review Committee, of the family and household structures, and taking the other measures set out in **Recommendation 5.2**.

24. PIA Topic 9: Expanding the use of PLIDA to support statistical production (including the Census and ABS Household Surveys) ³³

Description and information flows

- 24.1 From time to time, proposals have been considered, on a case by case basis, to use PLIDA data to support statistical production of particular data.
- 24.2 For example, separate PIA processes have been, or are currently being, considered for the use of PLIDA data:
- 24.2.1 to improve the efficiency of collection and quality of data collected by the ABS as part of the Census (e.g. to use information that indicates whether a property was not occupied on Census night, so that time and resources are not wasted in following up occupants to complete a Census form, thereby making the collection process more efficient); or
 - 24.2.2 to improve data quality and save survey respondents time and effort in completing surveys, by supplementing survey data with data items already included in PLIDA (e.g. to use information about Pharmaceutical Benefits Scheme (PBS) medications dispensed to a person to supplement their answers to the National Health Survey questions; to use PLIDA data to fill in missing information in the Census record or to add additional variables for questions that were not asked). ³⁴
- 24.3 The ABS wishes to consider more broadly the information flows and privacy risks associated with some key use cases, for which regular requests are made for PLIDA data to be used to support statistical production of ABS surveys (including the Census), specifically the use of PLIDA data for:
- 24.3.1 data supplementation – the use of PLIDA administrative data (i.e. data received from a government agency as the Data Custodian) to add more data items to other data which is collected as part of an ABS survey;
 - 24.3.2 data substitution – the replacement of survey data with administrative data contained in PLIDA (e.g. information in the survey about the individual's healthcare, income or tax, or childcare costs can be replaced with data in the PLIDA data which has already been collected);
 - 24.3.3 data editing – the use of PLIDA data to detect and correct errors in individual surveys; and
 - 24.3.4 survey process improvement – using PLIDA data to allow the ABS to be more efficient in conducting its survey operations. ³⁵
- 24.4 All of the usual PLIDA data governance processes would apply to proposals to use PLIDA data for any of these types of projects (i.e. the survey data would be added by the Data Custodian to PLIDA, and the Spine used to link survey data to other PLIDA datasets).

³³ In the Stakeholder Consultation Report, this topic is described as 'PIA Topic 5'.

³⁴ As these issues were, or are being, considered through separate PIA processes, they are not within the scope of the 2024-25 PIA Update.

³⁵ More information available here:

<https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Basic+Survey+Design+-+Data+Processing>

Discussion and analysis of relevant key privacy issues

Use of PLIDA data (APP 6)

- 24.5 All of the existing mitigation strategies which are used to reduce reidentification risks for data subjects for any approved research project, will apply to all of the use cases outlined above. We expect that these will continue to be appropriate to minimise compliance risks associated with this topic (that is, we do not consider that the use of PLIDA data for statistical production has a greater risk profile than any other approved research project, if normal PLIDA processes are followed – including to ensure that analytical and linkage data remain separated, with access by Authorised Users only provided to the analytical data).
- 24.6 I ABS will have collected the PLIDA data for the purposes of establishing and operating PLIDA through facilitating approved research projects. Given that the use cases will be approved PLIDA research projects, use of this data arguably falls within the primary purpose for collection by the ABS, consistent with APP 6.
- 24.7 More importantly, because PLIDA data used for the use case research projects will be analytical data only, personal information will not be analysed by Authorised Users (so there will be no breach of APP 6 in disclosing that data to them).
- 24.8 However, it is possible that an Authorised User, if employed or engaged by the Data Custodian that provided the original survey data to PLIDA, may have access to the original identified survey information in addition to linked analytical data (including deidentified survey data) via PLIDA. Such individuals (particularly any who are 'motivated intruders') may be in a position to potentially use the deidentified integrated PLIDA data, to reidentify the data subjects using the original survey data.
- 24.9 We understand that such a risk would be identified through usual PLIDA project approval processes, and note that:
- 24.9.1 those proposing the research project would need to explain how this risk would be managed (e.g. by ensuring that Authorised Users with access to the integrated PLIDA data did not also have access to the original survey data); and
- 24.9.2 the research project would not be approved if all Data Custodians whose data would be used for the project were not satisfied that the reidentification risk had been properly addressed.

Notification to data subjects (APP 5, APP 1)

- 24.10 There may be concerns that the survey respondents may have no awareness that other data that they have provided to another entity, such as a government agency, will be used to supplement or support that survey.
- 24.11 We therefore suggest that the ABS:
- 24.11.1 work with Data Custodians (including relevant parts of the ABS that conduct surveys, including for the Census and ABS Household Surveys) to ensure that their survey processes are designed to increase the likelihood that data subjects will understand how their survey answers may be supplemented by, or used in connection with, other data that they have provided to other entities (particularly to government agencies); and
- 24.11.2 continue to publish information about approved research projects, including examples of the use cases above, as an openness and transparency measure.

(Recommendation 6).

Proportionality (All APPs)

- 24.12 For all use cases it will be important that the nature of the particular survey, and the PLIDA data used to support or supplement it, will continue to be considered as part of the research project proposal. We note that the usual PLIDA processes for approving proposed research projects will include consideration of the specific details of a proposed project, including whether any privacy risk is proportionate to the anticipated benefits.
- 24.13 We note that, if there are no direct impacts on the individuals participating in the survey, the privacy impacts will be significantly reduced.

25. PIA Topic 10: Publication of aggregated PLIDA data

Description and information flows

- 25.1 Currently the ABS publishes aggregated data on its website when there is a clear benefit in informing the public, and there is a legal mechanism for its disclosure. For example, the ABS recently published the 'Administrative Data Snapshot' derived from PLIDA analytical data in the form of aggregated outputs about population and housing, in accordance with the Census and Statistics Act.
- 25.2 The ABS is considering whether to repeat this type of product and more generally publish outputs that contain integrated analytical PLIDA data on its website.³⁶ Such outputs would only contain PLIDA analytical data that has been aggregated by Statistical Area Level 2 (SA2)³⁷ or postcode (using the ABS Data by Region and/or the Digital Atlas of Australia). SA2 generally represents a community that interacts together socially and economically, with an average population of 10,000 people.
- 25.3 As was done for the Administrative Data Snapshot, the ABS intends that such outputs would only be created (and published) as part of an approved PLIDA project (i.e. all approvals required under the PLIDA governance framework, including from relevant Data Custodians, would apply), and the ABS' other internal processes for release of data would also apply.

Discussion and analysis of relevant key privacy issues

Reidentification risks (All APPs)

- 25.4 As some stakeholders raised in the consultation sessions, any published data will no longer protected by the PLIDA data governance framework, thus is at risk of being aggregated with other data available to the world at large, particularly noting there is currently no obligation on any member of the public accessing PLIDA data to not reidentify data subjects in deidentified data (unlike for Authorised Users who are required to agree to such an obligation).
- 25.5 Publication of *any* data will attract some level of 'motivated intruder risk' (that is, there are likely to be individuals who seek to access and misuse any published information, such as those using data-scraping technology that will collate ABS data with other publicly available data).

³⁶ The ABS is also considering this issue in the context of the Business Longitudinal Analysis Data Environment (BLADE), but this is out of scope for the 2024-2025 PIA Update.

³⁷ For more information about SA2, see <https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/jul2021-jun2026/main-structure-and-greater-capital-city-statistical-areas/statistical-area-level-2>.

- 25.6 The means that it is inherent for the ABS to pay particular attention to the data it seeks to aggregate and publish, to ensure that the reidentification risk can be validly assessed as 'very low'.
- 25.7 We understand that, particularly given the secrecy provisions in the Census and Statistics Act, the ABS has robust consideration and approval processes (including oversight by the ABS' Disclosure Review Committee) that must be followed before data held by the ABS can be published.
- 25.8 In particular, we understand that the ABS would conduct its usual assessments to be able to conclude that published data at the level of granularity of SA2 or postcodes strikes a sufficient balance between risk of reidentification and public utility of such data, such that the benefits from publication of the data is proportionate to the potential privacy risks involved.
- 25.9 On that basis, we do not consider that additional mitigation strategies need to be employed by the ABS before publication of aggregated, and appropriately deidentified, integrated PLIDA data.

Part H Glossary

25.10 In this report, the following terms and acronyms have the meanings in the table below.

Definitions	
2018 PIA	means the original MADIP PIA undertaken by the ABS in 2018, which assessed the privacy impacts associated with the establishment of the MADIP infrastructure.
2019 PIA Update	means the update of the MADIP PIA undertaken by the ABS in 2019, which considered planned changes or updates to PLIDA since 2018.
2022 PIA Update	means the update of the MADIP PIA undertaken by the ABS in 2022, which considered planned changes or updates to PLIDA since 2019.
ABS	means the Australian Bureau of Statistics.
ACT	means the Australian Capital Territory.
administrative data	means datasets that were originally collected by a Commonwealth or State or Territory government agency as part of their functions and activities, before the agency as Data Custodian provided that data for inclusion in PLIDA.
AIHW	means the Australian Institute of Health and Welfare.
analytical data	means other information about an individual that is the subject of a dataset which does not identify them, separate to 'linkage information'.
ANDII	means the underlying infrastructure to deliver the disability data asset. The ANDII refers to the: <ul style="list-style-type: none"> • national linkage and integration infrastructure, comprising a national spine and linkage model; and • national data integration system governance and streamlined data sharing arrangements that enable the creation of domain-specific data assets such as the National Disability Data Asset; and • Australian National Data Integration Infrastructure ICT Solution.
ANDII ICT solution	means the ICT environment that is part of ANDII which is currently used for processing and storing NDDA data in accordance with the NDDA governance arrangements.
APP, or Australian Privacy Principle	has the meaning given to it in the Privacy Act.
APP Code	means the <i>Privacy (Australian Government Agencies – Governance) APP Code 2017</i> .

Definitions	
APP Guidelines	means the <i>Australian Privacy Principles guidelines</i> published by the OAIC, found at https://www.oaic.gov.au/privacy/australian-privacy-principles/australian-privacy-principles-guidelines .
ARID	means address register ID number.
ABS	means the Australian Bureau of Statistics.
ASD	means the Australian Signals Directorate.
Authorised Users	means the users who have been authorised to access PLIDA data for approved research projects.
Census and Statistics Act	means the <i>Census and Statistics Act 1905</i> (Cth).
Core Demographics Module	means the PLIDA module using existing PLIDA data, which contains demographic items derived from data across the PLIDA datasets such as data subjects' age, gender, death dates, country of birth and highest level of education attained.
Core Indigenous Module	means the PLIDA module using existing PLIDA Data, which replaces the 'ever identified' derived Indigenous status contained in the Combined Demographics – Indigenous Module, and contains information about whether or not a data subject has identified as Aboriginal and/or Torres Strait Islander.
Core Locations Module	means the PLIDA module using existing PLIDA data, which contains existing geographic information about data subjects within PLIDA using information from administrative data within PLIDA.
Core Module	means the ensuring PLIDA dataset created from other PLIDA data, which can be used by Authorised Users for an approved PLIDA project. Core Modules include the Core Demographics Module, Core Indigenous Module, Core Locations Module, Core Scoping Module and Core Relationships Module.
Core Scoping Module	means the PLIDA module using existing PLIDA data, which collates information in PLIDA about data subjects' residence and activity in Australia.
Core Relationships Module	means the PLIDA module using existing PLIDA data, which will support research projects examining family and household analysis.
DAT Act	means the <i>Data Availability and Transparency Act 2022</i> (Cth).
Data Custodians	means the entities which are responsible for the preparation and provision of source data and metadata that is shared with and supplied to the ABS for inclusion in PLIDA, from whom approval is sought for each research project seeking to use that data.
DataLab	means the ABS DataLab, the cloud-based analysis solution which enables PLIDA Authorised Users to undertake real-time complex analysis of detailed analytical microdata.

Definitions	
DOMINO	means Data Over Multiple Individual Occurrences data asset, which contains snapshots of the characteristics of recipients of government payments such as the Age Pension and JobSeeker.
family structure	means the collection of data values created in connection with the LCDI, designed to identify data subjects sharing an address and how these individuals change over time.
household structure	means the collection of data values created in connection with the LCDI, designed to define relationships between a child data subject (i.e., a person aged 14 or under) and other data subjects who are their family members (e.g., their mother) and/or live in the same household with them (e.g., grandparents).
ISM	means the Australian Government Information Security Manual.
inherently sensitive survey data	means information that does not meet the legislative definition for 'sensitive information' in the Privacy Act, but nevertheless has an inherent degree of sensitivity (for example, information about an individual's experience with family, domestic or sexual violence).
LCDI	means the Life Course Data Initiative, a pilot project that will run from FY2023-24 to FY2026-27. This project involves adding new datasets to PLIDA from administrative data sources, and then using that and other PLIDA data in research projects to inform long-term policy responses to disadvantage.
linkage information	contains identifying information about each individual that is the subject of the dataset, separate to 'analytical data'.
IRAP	means the Australian Government Information Security Manual, one of the security arrangements in place for the IT systems used for PLIDA.
MADIP	means the Multi-Agency Data Integration Project, which is the name by which PLIDA was formerly known.
NDDA	means the data asset known as the 'National Disability Data Asset' which brings together deidentified information from different government agencies, to facilitate research projects to better understand the experiences of people with disability.
NDDA PIA	means the PIA process undertaken in relation to the NDDA and ANDII in 2022-23.
NFP organisations	means organisations who operate in a 'not for profit' manner, that is, they do not operate for the profit, personal gain or other benefit of particular people.
non-sensitive survey data	information that does not otherwise fall into the categories of sensitive survey data or inherently sensitive survey data.
OAIC	means the Office of the Australian Information Commissioner.
personal information	has the meaning given in section 6 of the Privacy Act.

Definitions	
PIA	means a privacy impact assessment.
PIA Topic	means a topic that is within the scope of this PIA Update process, as described in this report.
PIA Update	means this PIA process in respect of PLIDA, which is designed to update the previous 2018 PIA, 2019 PIA Update, and the 2022 PIA Update, in relation to the PIA Topics.
PLIDA	means Person Level Integrated Data Asset, a secure, person-based data asset that combines a broad set of information on health, education, government payments, income and taxation, employment, and population demographics to create a comprehensive dataset to facilitate the use and reuse of public data for research purposes.
PMP	means the PLIDA Modular Product.
Privacy Act	means the <i>Privacy Act 1988</i> (Cth).
PTA	means a privacy threshold assessment, which is a mechanism to determine whether a project represents a 'high privacy risk'.
SA2	means Statistical Area Level 2.
SDIE	means the Secure Data Integration Environment, the current infrastructure in which PLIDA data is stored and processed.
sensitive information	has the meaning given in section 6 of the Privacy Act.
Spine	means the Person Linkage Spine, which is the central linkage infrastructure within PLIDA that creates the capacity for separate source datasets to be linked.
sensitive survey data	is a subset of survey data, being data that would (if the data subject was identified), falls within one of the categories set out in the definition of 'sensitive information' under the Privacy Act, such as personal information about an identified individual's racial or ethnic origin, sexual orientation, criminal record or health, or their genetic or biometric information.
survey data	means data collected by a Data Custodian from data subjects via a survey.

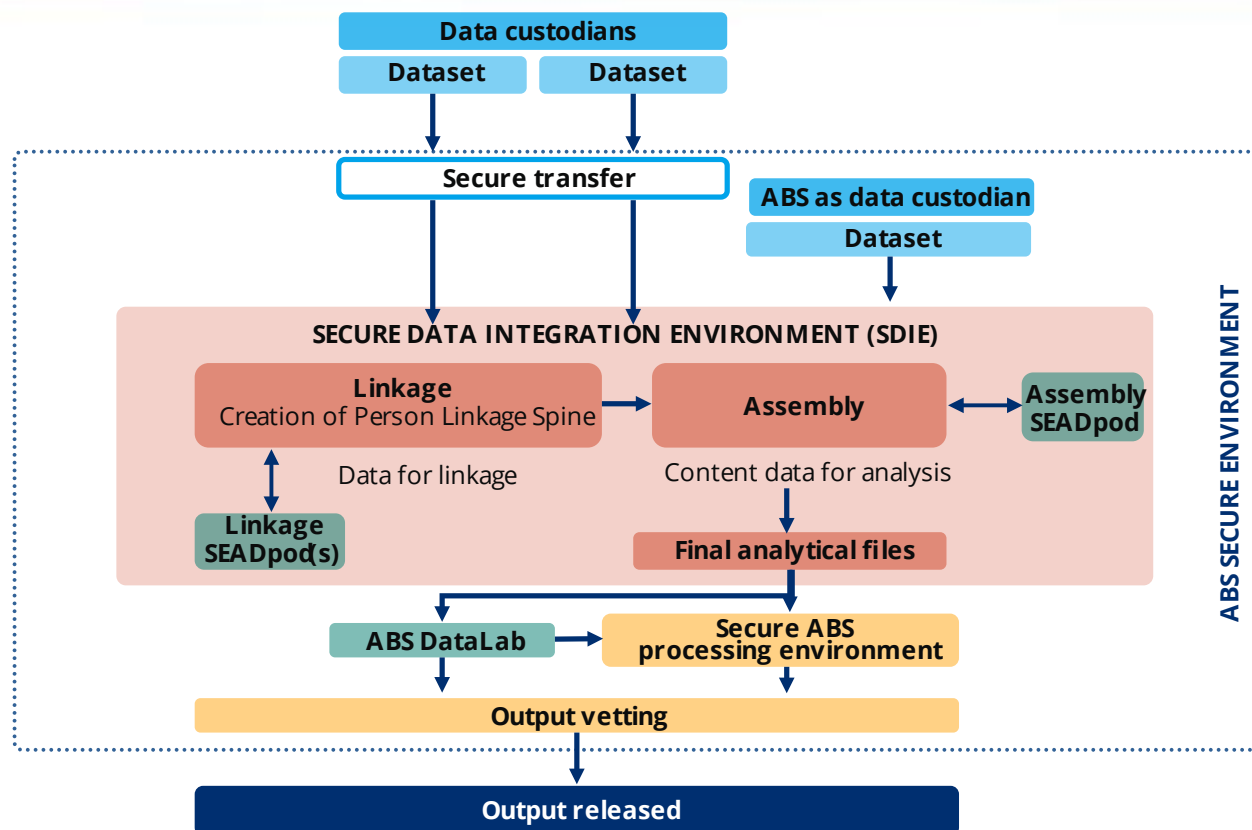
Attachment 1 Materials Reviewed

This Attachment sets out a list of materials provided to Maddocks in connection with our conduct of this PIA. It does not include additional publicly available materials, including legislation, guidance and research materials, that we also considered as part of this PIA.

- PLIDA PIA_ Document Register_29.08.2024 (received by Maddocks on 9 September 2024);
- 1. PLIDA PIA_Scope_09.09.2024 (received by Maddocks on 9 September 2024);
- PLIDA PIA Consultation Plan_29.08.2024 (received by Maddocks on 9 September 2024);
- PTA Creation and Use of Five New PLIDA Modular Products Endorsed 15072024 (received by Maddocks on 9 September 2024);
- PTA High-level administrative crime and justice data linkage to PLIDA Endorsed 15082024 (received by Maddocks on 9 September 2024);
- PTA Re-use of PLIDA data for NDDA_ Privacy Threshold Assessment Endorsed (received by Maddocks on 9 September 2024);
- LCDI household and families information (received by Maddocks on 9 September 2024);
- [From ABS 1.10.24] Simple PLIDA data flow (final version) (received by Maddocks on 1 October 2024); and
- 2. [Rec'd ABS 1.10.24] Difference between core relationship and LCDI modules (received by Maddocks on 1 October 2024).

Attachment 2 PLIDA Information Flow Diagram – Current Data Handling Practice

PLIDA information flow diagram – current data handling practice

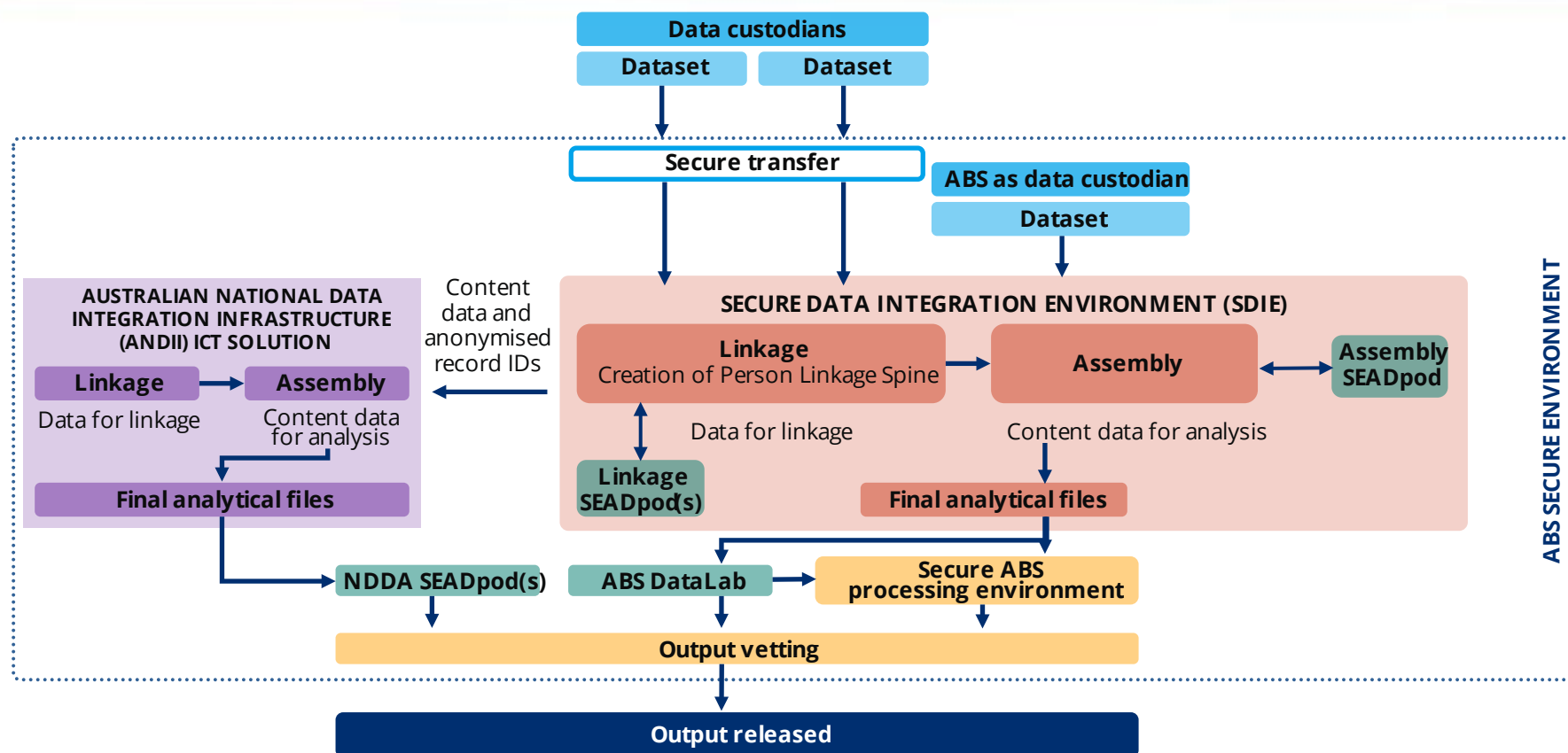


1

This diagram was generated and provided by the ABS.

Attachment 3 PLIDA Information Flow Diagram – Proposed Data Handling Practice

PLIDA information flow diagram – proposed data handling practice



2

This diagram was generated and provided by the ABS.

Attachment 4 PLIDA datasets

Datasets in PLIDA – Jurisdiction, Data custodian and Type of linkage by Dataset name and Linkage information

Datasets currently linked to PLIDA (formerly referred to as MADIP)

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Australian Bureau of Statistics						
Enduring linkages						
ABS Business Characteristics Survey	2005-06 to 2017-18	An annual survey providing estimates in business use of information technology, innovation, and a broad range of other non-financial business characteristics.	Census and Statistics Act 1905	N		N
ABS National Health Survey	2014-15, 2017-18, 2020-21	Information about the health of Australians, including prevalence of long-term health conditions health risk factors such as smoking, overweight & obesity, alcohol consumption & physical activity & demographic & socioeconomic characteristics.	Census and Statistics Act 1905	N	Disability	Y

³⁸ Includes the following populations: Children, Aboriginal and Torres Strait Islander, Disability, Aged.

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
ABS Survey of Disability, Ageing & Carers	2018	Data on people with disability, older people (aged 65 years or more) & people who care for people with disability or older people	Census and Statistics Act 1905	N	Disability, Aged	Y
Australian Census Longitudinal Dataset	2006, 2011, 2016	Brings together a five per cent sample from the 2006 Census with corresponding records from the 2011 & 2016 Censuses	Census and Statistics Act 1905	N	Children, Aboriginal and Torres Strait Islander, Disability, Aged	Y
Business Longitudinal Analysis Data Environment (BLADE) – CORE Dataset: Indicative data items; BLADE Locations data; Business Activity Statement; Business Income Tax; Payment Summaries Pay As You Go	From 2001	BLADE is an economic data tool combining tax, trade, and intellectual property information with ABS data to provide a better understanding of the Australian economy and business performance over time. A limited set of BLADE information (generally from 2001 onwards) is available in PLIDA.	Census and Statistics Act 1905	N		N
Census of Population and Housing (Census)	2011, 2016, 2021	Demographic information such as family composition, education attainment, marital status, and household income	Census and Statistics Act 1905	N	Children, Aboriginal and Torres Strait Islander, Aged, Disability	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Once-off linkages						
ABS Household Income & Expenditure Survey	2015-16	Two phase survey of Survey of Income & Housing and the Household Expenditure Survey which collect information on sources of income, amounts received, household net worth, housing, household characteristics & personal characteristics.	Census and Statistics Act 1905	N	Disability	Y
ABS Labour Force Survey	2020-2024	Information on how people manage their time and includes information such as indigenous status, marital status and employment status	Census and Statistics Act 1905	Y	Children, Aboriginal and Torres Strait Islander, Aged, Disability	Y
ABS Survey of Income and Housing	2019-20, 2020-21, 2022-2023	Information on sources of income, amounts received, household net worth, housing, household characteristics and personal characteristics collected from a sample of private households.	Census and Statistics Act 1905	N		N
Census ADVentures Housing Dataset	2017 - 2023	This dataset includes solar capacity and installation date, dwelling code and remoteness code, as well as the Census Enumeration Management Information (EMI)	Census and Statistics Act 1906	Y		N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Census Post Enumeration Survey	2021	Census Post Enumeration Survey is linked for experimental purposes. This data is not made available to external researchers*	Census and Statistics Act 1906	Y	N/A	Y
Australian Children's Education and Care Quality Authority (ACECQA)						
Once-off linkages						
National Quality Standard	From 2011	The NQS is a measure for early childhood education and care and outside school hours care services in Australia. Services are assessed and rated by ACECQA and given a rating for each of the 7 quality areas and an overall rating based on these results.	Education and Care Services National Law s271 (1c)	N		N
Australian Institute of Health and Welfare (AIHW)						
Once-off linkages						
National Disability Data Asset (NDDA) Pilot dataset extract	2010 to 2020	The NDDA is a joint integrated data project aiming to bring together Commonwealth and state and territory data to provide new insights for the future development of policy for people with disability.	Project approval by AIHW Ethics Committee (reference EO2020/3/1185), pursuant to Sections 5 and 6 and consistent with Section 29 of the Australian Institute of Health and Welfare Act 1987 (Cth). Collected under the authority of the Census and Statistics Act 1905 (Cth).	N	Children, Disability, Aboriginal and Torres Strait Islander	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Australian Taxation Office (ATO)						
Enduring linkages						
JobKeeper	2020 to 2021	Information about JobKeeper subsidy payments to businesses affected by coronavirus (COVID-19).	ATO data is disclosed to the Australian Statistician under the Taxation Administration Act 1953 (Cth) for the 'purpose of administering the Census and Statistics Act 1905' under which the MADIP is conducted. The Tax Law Amendment (Confidentiality of Taxpayer Information) Act 2010 (Cth) enables the ATO to provide the ABS with unit record PIT data.	N		N
Personal Income Tax (PIT) data: ATO Client Register, PAYG Payment Summaries, Client Data Income Tax Returns	From 2006 (Client Register), From 2010 (Payment Summaries and Income Tax Returns)	PIT comprises of: Client Register: demographic data about individuals who require a tax file number to interact with government, business, financial, educational and other community institutions; Pay as you go (PAYG) Payment Summaries: employer-issued records of payments made to individuals; Client Data Income Tax Returns: the tax return data filed by individuals	Taxation Administration Act 1953 for the 'purpose of administering the Census and Statistics Act 1905' – under which the MADIP is conducted. The Tax Law Amendment (Confidentiality of Taxpayer Information) Act 2010 enables the Australian Taxation Office to provide the ABS with unit record data.	N		Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Single Touch Payroll	From 2020-21	Information about employees' salaries and wages, pay as you go (PAYG) withholding and superannuation as reported through the Single Touch Payroll system.	Within the ATO, taxation data is collected under the Taxation Administration Act, 1953 (Cth), which allows the ATO to provide the Australian Statistician with this information for the purpose of administering the Census and Statistics Act, 1905 (Cth) – under which this product is created. The ABS acquires all ATO data, including STP job holder data, under the Census and Statistics Act, 1905 (Cth).	N		Y
Once-off linkages						
Early Release of Superannuation (ERS)	Apr 2020 to Dec 2020	Information about early access to retirement savings by eligible individuals.	ATO data is disclosed to the Australian Statistician under the Taxation Administration Act 1953 (Cth) for the 'purpose of administering the Census and Statistics Act 1905' under which the MADIP is conducted. Section 355-65 of Schedule 1 to the TAA contains specific exceptions about disclosing protected information for other government purposes which covers ERS data.	N		N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Job Maker Hiring Credit Dataset	1 Feb 2021 to 31 January 2023	Employer and employee information provided by businesses claiming a JobMaker Hiring Credit (JMHC). The JMHC scheme is an incentive for businesses to employ additional young job seekers aged 16–35 years.	ATO data is disclosed to the Australian Statistician under the Taxation Administration Act 1953 (Cth) for the ‘purpose of administering the Census and Statistics Act 1905’ under which the MADIP is conducted. The Tax Law Amendment (Confidentiality of Taxpayer Information) Act 2010 (Cth) enables the ATO to provide the ABS with unit record PIT data.	N	N	N
Member Account Transaction Service (MATS) and Member Account Attributes Service (MAAS) – Superannuation Balances data	2018-19, 2019-20	Information about superannuation contributions, transactions, and events, such as account closure and account phase. This information is reported to the ATO by regulated funds using the ATO’s MATS and MAAS services.	ATO data is disclosed to the Australian Statistician under the Taxation Administration Act 1953 (Cth) for the ‘purpose of administering the Census and Statistics Act 1905’ under which the MADIP is conducted. Section 355-65 of Schedule 1 to the TAA contains specific exceptions about disclosing protected information for other government purposes.	N		N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Taxable Payment Annual Report (TPAR) from Residential Aged Care Facilities, Payee Data	2020-21	Total payments made to contractors by residential aged care facilities, including subcontractors, consultants and independent contractors operating as companies, partnerships, trusts or sole traders.	ATO data is disclosed to the Australian Statistician under the Taxation Administration Act 1953 (Cth) for the 'purpose of administering the Census and Statistics Act 1905' under which the MADIP is conducted.	N	N	N
Private Health Insurance (PHI) Register	2012/13 – 2021/22	The Private Health Insurance Report is provided to the ATO from individual insurers on an annual basis. This report is the most complete snapshot of individuals that hold a private health insurance product.	Taxation Administration Act 1953 Schedule 1, Section 355-65(8), Table 7(1) Disclosure of protected information can be made to the Australian Statistician where the disclosure is for the purpose of the Statistician administering the Census and Statistics Act.	Y	N	Y
Department of Defence and Department of Veterans' Affairs						
Once-off linkages						
Veterans Population Spine and Analytical Dataset	1985 to 2020	Demographic, health, and service record information for members of the Australian Defence Forces (ADF), integrated by the Australian Institute of Health and Welfare (AIHW).	The components of this integrated dataset were collected from the data custodians by the Australian Institute of Health and Welfare pursuant to Section 5 Australian Institute Health and Welfare Act 1987 (Cth).	N	N	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
			The AIHW Ethics Committee granted a Waiver of Consent with respect to what would otherwise comprise breaches of APPs 3 and 6 pursuant to s.95 for the Privacy Act 1988(Cth) regarding disclosure / release of data from Departments of Defence and Department of Veterans Affairs to the AIHW.			
Department of Education						
Enduring linkages						
Australian Early Development Census (AEDC)	2009, 2012, 2015, 2018, 2021	A nationwide triennial census that looks at children in their first year of full-time school and measures how well children are developing across five important domains using an Early Development Instrument (EDI)	No specific legislation applied to the Australian Early Development Census, so the provisions of the Privacy Act 1988 are applied. A Memorandum of Understanding is in place. Information is shared in accordance with the Census and Statistics Act 1905 the Commonwealth Arrangements for Data Integration and the AEDC Data Guidelines.	N	Children, Aboriginal and Torres Strait Islander	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Higher Education Information Management System	From 2005	All higher education & Vocational Education and Training FEE-HELP data reported to the government. Domestic students enrolled in higher education on Commonwealth supported places	Higher Education Support Act 2003 & the VET Student Loans Act 2016	N	Children, Aboriginal and Torres Strait Islander, Disability	Y
Total VET Activity (TVA)	From 2015	Information about students and the Vocational Education and Training (VET) delivery activities of registered training organisations.	TVA analytical data is disclosed under Section 210A of the National Vocational Education and Training Regulator Act 2011 (Cth) and collected under the authority of the Census and Statistics Act 1905 (Cth). Linkage data is disclosed under the authority of subsection 22(5) of the Student Identifiers Act 2014 (Cth) and collected under the authority of the Census and Statistics Act 1905 (Cth).	N	Children, Aboriginal and Torres Strait Islander, Disability	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Once-off linkages						
Child Care Management System (data and linkage file)	2010 to 2018	Administrative data covering enrolment & attendance of children aged 4-6 [inclusive] & their associated carers, including basic demographics. This information is reported to the Department of Education, Skills and Employment and used to calculate the childcare fee reductions to be paid to the service.	Data is disclosed to the ABS pursuant to a public interest certificate (PIC) issued in according with paragraph 168(1)(a) of the A New Tax System (Family Assistance) (Administration) Act 1999	N	Children, Aboriginal and Torres Strait Islander	Y
Employment Services System (ESS) and Post-Program Monitoring Surveys (PPM)	2009-current	Demographic and employment outcomes information collected from Job Seekers during the delivery of employment services.	Social Security (Administration) Act 1999, subsection 202(2C)	N	Aboriginal and Torres Strait Islander, Disability	Y
NSW Family Day Care data	2013 to 2018	Information on children participating in early childhood education in NSW, sourced from the Child Care Subsidy System data.	A New Tax System (Family Assistance) (Administration) Act 1999	N	Children	Y
NSW Long Day Care submission for National Early Childhood Education and Care Collection	2013 to 2018	Information on children participating in early childhood education in NSW.	Previously supplied for an ABS publication and is covered by the Census and Statistics Act 1905.	N	Children	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Student Residential Address and Other Information (Student) collection	From 2018	Information about student residential address (excluding names), level of schooling, boarding status and names and addressed of students' parents and/or guardians, from eligible schools. The Student Residential Address and Other Information Collection (Address Collection) is undertaken each year by non-government Approved Authorities and their schools.	Australian Education Act 2013 section 125(1), Australian Education Regulation 2013 section 21, section 58B, and section 65.	N	Children	N
Provider Registration and International Students Management System (PRISMS) (from 2006)	From 2006 to latest available	Information on the course enrolments of overseas students.	Census and Statistics Act 1905	N		Y
Child Care Administrative data	From 2019-2021	CCA data contains both Child Care Management System (CCMS) and Child Care Subsidy System (CCSS) data.	Census and Statistics Act 1905	N	Children	Y
Department of Employment and Workplace Relations						
Enduring linkages						
Australian Apprenticeship and Incentives Program (AAIP) and Training Contracts	From 2006-2023	Information about trainees & apprentices, qualifications, their employers/trainers & incentive payments provided through the program	MoU between the ABS and DEWR	N	Aboriginal and Torres Strait Islander, Disability, Children	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Once-off linkages						
Department of Health and Aged Care						
Enduring linkages						
Australian Immunisation Register	From 2010	Information about COVID-19 and other vaccination status, including core demographics	Australian Immunisation Register Act 2015 (Cth) section 6 and section 9 and pursuant to a Public Interest Authorisation issued by the Minister under section 22(3).	N	Children, Aboriginal and Torres Strait Islander, Aged	Y
Medicare Benefits Schedule (MBS)	From 2005	Information on the usage of Medicare-subsidised health care services, such as General Practitioner attendances, mental health, and pathology	Data is disclosed to the ABS pursuant to public interest certificates (PICs) issued the Minister (or delegate) in accordance with section 130 of the Health Insurance Act 1973 and section 135A of the National Health Act 1953. In accordance with the National Health Act 1953, Medicare Benefits Schedule and Pharmaceutical Benefits Scheme source data is acquired and stored separately in the MADIP.	N	Children, Disability, Aged	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Pharmaceutical Benefits Schedule (PBS)	From 2006	Information about use of prescription medications and services subsidised under the Pharmaceutical Benefits Scheme	Data is disclosed to the ABS pursuant to public interest certificates (PICs) issues by the Minister (or delegate) in accordance with section 130 of the Health Insurance Act 1973 and section 135A of the National Health Act 1953. In accordance with the National Health Act 1953, Medicare Benefits Scheme and Pharmaceutical Benefits Schedule source data is acquired and stored separately in the MADIP.	N	Children, Disability, Aged	Y
Once-off linkages						
Aged Care Provider data	From 2015	Information on Aged Care Homes, including permanent and respite residential care and Multi-Purpose Service residential care, sourced from the public dataset at GEN Aged Care Data.	Public dataset – see AIHW Act 1987 Section 6(f)	N	Aged	N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Centralised Register of Medical Practitioners (Provider Directory)	From 2011	Information about registered medical practitioners, including specialities	Data is disclosed to the ABS pursuant to public interest certificates (PICs) issued by the Minister (delegate) in accordance with section 130 of the Health Insurance Act 1972 and section 135A of the National Health Act 1953	N	Children, Aged, Disability	Y
Derived Medical Speciality	From 2014/15 – 2023/24	Derived Medical Specialty (DMS) data for calendar years 2014 - 2023 and financial years 2014-15 to 2023-24. The population in scope is all Medical and Allied Health practitioners who billed Medicare.	Health Insurance Act 1973, Census and Statistics Act 1905	Y	N	N
Department of Home Affairs						
Enduring linkages						
Migration data – Adult Migrant English Program data	From 2003-2019	Information on participants in the English language tuition program, including their visa type, course attendance and educational outcomes.	Census and Statistics Act 1905	N		N
Migration data – Client information	From 1984	Client information on Australian-born citizens, temporary & permanent migrants	Migration Act 1958 & the Australian Border Force Act 2015	N		Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Migration data – Skilled Migration Points	From 2005	Information on the points assigned across 30 fields related to skills & experience for persons who have applied for skilled migration visas	Migration Act 1958 & the Australian Border Force Act 2015	N		Y
Migration data – Traveller data	From 2004	All overseas movement records on Home Affairs' Travel & Immigration Processing System (TRIPS). Movement records are supplied via monthly extracts with the data being compiled on a quarterly basis	Migration Act 1958 & the Australian Border Force Act 2015	N		Y
Migration data – Visa information & Citizenship grants	From 2000	Information on visa types, start & end dates, & data on educational studies for student visas & working arrangements for skilled migration visas	Migration Act 1958 & the Australian Border Force Act 2015	N		Y
Department of Social Services						
Enduring linkages						
Data Exchange (DEX) Dataset	From July 2015	Data from the program performance reporting tool that allows funded organizations to report their service delivery information & demonstrate outcomes.	S208 of the Social Security (Administration) Act 1999	N	Disability, Aboriginal and Torres Strait Islander	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
DOMINO Centrelink Administrative data (formerly Social Security and Related Information (SSRI))	From 2006	Data Over Multiple Individual Occurrences (DOMINO) contains snapshots of the characteristics of recipients of government payments such as the Age Pension and JobSeeker (formerly Newstart Allowance). This dataset was previously known as Social Security and Related Information (SSRI).	Data is provided pursuant to PICs issued under the following provisions which allow for the disclosure of personal and sensitive information where it is in the public interest: • S208 of the Social Security (Administration) Act 1999; • S168 of the A New Tax System (Family Assistance) (Administration) Act 1999; • S128 of the Paid Parental Leave Act 2010; and • S355 of the Student Assistance Act 1973.	N	Aboriginal and Torres Strait Islander, Disability, Aged	Y
Once-off linkages						
DES Participant Dataset DES Service Placement Dataset DES Job Placement Dataset	From 2018-2022	These datasets were a once-off linkage including data pertaining to participant interventions, job vacancies and use of social services.	Data is disclosed to the ABS pursuant to a public interest certificate (PIC) issued in according with paragraph 168(1)(a) of the <i>A New Tax System (Family Assistance) (Administration) Act 1999</i> .	Y	Aboriginal and Torres Strait Islander, Disability	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
Transgenerational Data Set (TDS) extract	From 1993 to 2014	Links the social assistance records of a birth cohort of young Australians to their parents, including payment history, demographic information, and accommodation and study history.	•S202(2C) of the Social Security (Administration) Act 1999; •S168(1)(a) of the A New Tax System (Family Assistance) Administration Act 1999; •S355(1)(a) of the Student Assistance Act 1973; & s128 (1)(a) of the Paid Parental Leave Act 2010. • Public Interest Certificate supplied by the Department of Social Services. (Signed 10 July 2019)	N	Children, Aged	Y
Jobs and Skills Australia						
Once-off linkages						
Foundational and Pathways Courses / Microcredentials	2022	A list of nationally recognised Micro-credentials offered in Australian Vocational Educational Training Institutions.	Letter of Exchange	Y	N	N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
National Disability Insurance Agency						
Enduring linkages						
National Disability Insurance Scheme (NDIS) Participant and Provider Data	2013 to 2023	Information on primary disability, including services attained, by participants in the NDIS.	Data is disclosed to the ABS pursuant to a public interest certificate (PIC) issued in accordance with the NDIS (Protection and Disclosure of Information) Rules 2013. National Disability Insurance Scheme Act 2013 Part 2, Div 1 C60(3) and National Disability Insurance Scheme Act 2013 Part 2, Div 1 C66 (1).	N	Children, Disability	Y
Once-off linkages						
NDIS Data Plan	From 2013	The data includes participant demographic information, such as the participant's country of birth and Aboriginal and Torres Strait Islander status. There is also outcome data which includes information about the individual's relationships, self-reported health status and education.	Data is disclosed to the ABS pursuant to a public interest certificate (PIC) issued in accordance with the NDIS (Protection and Disclosure of Information) Rules 2013. National Disability Insurance Scheme Act 2013 Part 2, Div 1 C60(3) and National Disability Insurance Scheme Act 2013 Part 2, Div 1 C66 (1).	Y	Disability, Aboriginal and Torres Strait Islander, Aged	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION ³⁸	SENSITIVE DATA
NDIS Funding Data	From 2013	The data includes payment information such as participant's identifier and claims.	Data is disclosed to the ABS pursuant to a public interest certificate (PIC) issued in accordance with the NDIS (Protection and Disclosure of Information) Rules 2013. National Disability Insurance Scheme Act 2013 Part 2, Div 1 C60(3) and National Disability Insurance Scheme Act 2013 Part 2, Div 1 C66 (1).	Y	Disability, Aboriginal and Torres Strait Islander, Aged	Y
Services Australia						
Enduring linkages						
Medicare Consumer Directory	From 2006	Information on persons enrolled with Medicare. This dataset was previously known as the Medicare Enrolments Database.	Data is disclosed to the ABS pursuant to public interest certificates (PICs) issued the Minister (or delegate) in accordance with section 130 of the Health Insurance Act 1973 and section 135A of the National Health Act 1953.	N	Children, Aged	Y

STATE AND TERRITORY DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2024/25 PLIDA PIA UPDATE	VULNERABLE POPULATION	SENSITIVE DATA
State and Territory Registrars of Births, Deaths, and Marriages						
Enduring linkages						
Death Registrations	From 2005	Information relating to registered death records from Australian States and Territories.	Deaths data are supplied for use in the MADIP under the Census and Statistics Act 1905, pursuant to an MOU with the QLD Registrar for Births, Deaths, and Marriages (on behalf of the other State and Territory Registrars). The Registrars of Births, Deaths, and Marriages (RBDMs) are authorised to collect and provide access to this data by their respective Births, Deaths and Marriages Acts.	N	Aboriginal and Torres Strait Islander	Y
Once-off linkages						
Birth Registrations	From 2006 to 2023	Information collected by State and Territory Registrars to enable the registration of a birth.	Statistics (Arrangements with States) Act 1956 (5)(1, 2)	N	Aboriginal and Torres Strait Islander	Y
Cancer Alliance Queensland						
Once-off linkages						
Cancer Institute of New South Wales						
Once-off linkages						

STATE AND TERRITORY DATA CUSTODIANS						
NSW Linked Cancer dataset: BreastScreen NSW	From 1988 to 2020	Information on people presenting for breast cancer screening in NSW, including assessment and treatment information and family histories of cancer.	The New South Wales Health Administration Regulation 2015 allows for the disclosure of epidemiological data held by the Cancer Institute New South Wales on behalf of the New South Wales Ministry of Health that does not identify any individual to whom the information relates for statistical & research purposes.	N	Aboriginal and Torres Strait Islander, Aged	Y
NSW Linked Cancer dataset: Cancer Incidence and Mortality	From 1972 to 2018	Demographic, cancer diagnosis, and mortality information for people with cancer in NSW.	The New South Wales Health Administration Regulation 2015 allows for the disclosure of epidemiological data held by the Cancer Institute New South Wales on behalf of the New South Wales Ministry of Health that does not identify any individual to whom the information relates for statistical & research purposes.	N	Aboriginal and Torres Strait Islander, Aged	Y
CINSW Pap Test Registry	From 1996 to 2017	Information on women who have a pap test in NSW, including the results of cervical cytology, histology tests or HPV DNA tests.	The New South Wales Health Administration Regulation 2015 allows for the disclosure of epidemiological data held by the Cancer Institute New South Wales on behalf of the New South Wales Ministry of Health that does not identify any individual to whom the information relates for statistical & research purposes.	N	Aboriginal and Torres Strait Islander, Aged	Y

STATE AND TERRITORY DATA CUSTODIANS						
NSW Department of Communities and Justice						
Once-off linkages						
New South Wales Human Services Dataset	From 1990 to 2022	The HSDS brings together NSW and Commonwealth records about children, young people and families, from across government services. The service streams, outcomes and life events captured in the HSDS included child protection, justice, housing, health and mental health, education, parental risk indicators and alcohol and other drug use.	Public Interest Direction made under s41(1) of the Privacy and Personal Information Protection Act 1998, together with s19(2)(a) of the PPIP Act and the Health Public Interest Direction made under s62(1) of the Health Records and Information Privacy Act 2012 and ss 14(a) and 15(2)(a) of Schedule 1 of the HRIP Act 2002.	N	Children, Aboriginal and Torres Strait Islander	Y
NSW Department of Education						
Once-off linkages						
Annual (Community) Preschool Census	From 2016	The census collects information from NSW community preschools that feeds into the National Early Childhood Education and Care Census Collection (NECECC). The scope of the NECECC consists of all service providers delivering a preschool program to children aged 3 to 6 years (inclusive) enrolled during the reference period.	Memorandum of Understanding between NSW Department of Education and ABS (updated variation will be drafted mid-year following ethics approvals); ABS and CHeReL will have a Letter of Exchange and Confidentiality Undertaking	Y	Aboriginal and Torres Strait Islander, Children	Y

STATE AND TERRITORY DATA CUSTODIANS						
NSW Early Childhood Outcomes: Best Start Kindergarten	2015, 2018, 2019	Information on children participating in early childhood education in NSW.	NSW Privacy and Personal Information Protection Act 1998 s18(1)(a) and s19(2)(a)	N	Children	Y
Government school students Socio-Educational Advantages Indexes	From 2017-2021	Student-level information with SEA index attached.	NSW Research Exemption under section 27B of the Privacy and Personal Information Protection Act 1998 (NSW)	N	Children	N
Mid-Year Census – NSW Government Preschools and Early Intervention	From 2016	The census collects information from NSW community preschools that feeds into the National Early Childhood Education and Care Census Collection (NECECC). The scope of the NECECC consists of all service providers delivering a preschool program to children aged 3 to 6 years (inclusive) enrolled during the reference period.	Memorandum of Understanding between NSW Department of Education and ABS (updated variation will be drafted mid-year following ethics approvals); ABS and CHeReL will have a Letter of Exchange and Confidentiality Undertaking	Y	Children, Aboriginal and Torres Strait Islander	Y
My school master file	2016-2023	School-level data including location, school sector, type, enrolments and funding.	Census and Statistics Act 1905	Y	Children	N

STATE AND TERRITORY DATA CUSTODIANS						
Nationally Consistent Collection of Data on school students with disability (NCCD)	From 2015	The Nationally Consistent Collection of Data (NCCD) on students with disability is an annual data collection from all Australian schools. Data is based on teacher judgement and reflects adjustments made for students with disability as defined by the Disability Discrimination Act 1992 (DDA). It includes, but is not limited to, students with diagnosed disabilities.	Memorandum of Understanding between NSW Department of Education and ABS (updated variation anticipated to be drafted mid-year following ethics approvals); ABS and CHeReL will have a Letter of Exchange and Confidentiality Undertaking	Y	Children, Disability	Y
NSW Early Childhood Outcomes: Community and Government Preschool Data	2013, 2014, 2016 to 2018	Information on children participating in early childhood education in NSW.	NSW Privacy and Personal Information Protection Act 1998 s18(1)(a) and s19(2)(a)	N	Children	Y
New South Wales National Assessment Program - Literacy and Numeracy (NAPLAN) data, Year 3	2018	Outcomes from the National Assessment Program - Literacy and Numeracy (NAPLAN) tests for New South Wales students who were in Year 3 in 2018.	NSW Privacy and Personal Information Protection Act 1998 s18(1)(a) and s19(2)(a)	N	Children	Y
New South Wales public school enrolment and student characteristics	1996 to 2016	Student characteristics including enrolments and attendance, course data and field of education.	NSW Research Exemption under section 27B of the Privacy and Personal Information Protection Act 1998 (NSW)	N	Children	N

STATE AND TERRITORY DATA CUSTODIANS						
New South Wales public school workforce profile (teachers)	1996 to 2016	Information on school and teacher characteristics.	NSW Research Exemption under section 27B of the Privacy and Personal Information Protection Act 1998 (NSW)	N	Children	N
Pathways for the Future dataset	From 1996	Population ages 0-67, who attended secondary school, undertook a vocational education and training qualification or an apprenticeship, traineeship or higher education qualification in NSW	Memorandum of Understanding between NSW Department of Education and ABS; ABS and CHeReL have a Letter of Exchange and Confidentiality Undertaking Ethics Approval	Y	Children, Aged, Disability, Aboriginal and Torres Strait Islander	Y
Record of School Achievement (RoSA) / Higher School Certificate (HSC)	2000-2021	Student level with information around school enrolment and course results	NSW Research Exemption under section 27B of the Privacy and Personal Information Protection Act 1998 (NSW)	N	Children	N
School information	2016-2024	School-level data, including location, and derived scores of occupation, education and income, of families / parents	Census and Statistics Act 1905	Y	Children	N
School recurrent funding model	2018-2030	School-level data, including estimated Australian Government Funding.	Census and Statistics Act 1905	Y	Children	N

STATE AND TERRITORY DATA CUSTODIANS						
Student Address	2018-2024	These files contain school level data. They enable the analysis of funding implications resulting from changes in the school scores, as well as providing publicly available information about schools (such as location, demographic information, etc).	Census and Statistics Act 1905	Y	Children	N
NSW Department of Industry						
Once-off linkages						
New South Wales Smart, Skilled and Hired data	1996 to 2016	Characteristics and outcomes of young people engaged in the NSW Government Smart, Skilled and Hired youth employment program.	NSW Research Exemption under section 27B of the Privacy and Personal Information Protection Act 1998 (NSW)	N		
NSW Education Standards Authority						
Once-off linkages						
New South Wales National Assessment Program - Literacy and Numeracy (NAPLAN), Record of School Achievement, Higher School Certificate	1996 to 2021	Student and family characteristics, test results and other education outcomes.	NSW Research Exemption under section 27B of the Privacy and Personal Information Protection Act 1998 (NSW)	N	Children	

STATE AND TERRITORY DATA CUSTODIANS

NSW Ministry of Health

Once-off linkages

NSW Admitted Patient Data	From 2001	The NSW Admitted Patient Data Collection (APDC) records all inpatient separations (discharges, transfers and deaths) from all public hospitals, public psychiatric hospitals, multi-purpose services, private hospitals, and private day procedure centres in NSW. Approximately 400 facilities contribute to the data collection, which includes patient demographic information, diagnoses, procedures, and administrative information such as dates of admission and separation, source of referral to the service, service referred to on separation, and patient health insurance status.	disclosed by NSW Ministry of Health to the ABS under the authority of section 27B of the <i>Privacy and Personal Information Protection Act 1998</i> (NSW) and under section 10(1)(f) and 11(1)(f) of Schedule 1 of the <i>Health Privacy Principles in the Health Records and Information Privacy Act 2002</i> (NSW) and collected under the authority of the <i>Census and Statistics Act 1905</i> (Cth).	N	Indigenous	Y
---------------------------	-----------	--	---	---	------------	---

STATE AND TERRITORY DATA CUSTODIANS						
NSW Ambulance	From 2009 - 2022	all persons aged 25 years and over at the time of the 2016 census (9 August 2016) who have had a CVD event/procedure in any of the multiple diagnosis or procedure fields in NSW APDC, NSW EDDC or anyone with a confirmed case of COVID-19 (regardless of age).	disclosed by NSW Ministry of Health to the ABS under the authority of section 27B of the <i>Privacy and Personal Information Protection Act 1998</i> (NSW) and under section 10(1)(f) and 11(1)(f) of Schedule 1 of the <i>Health Privacy Principles in the Health Records and Information Privacy Act 2002</i> (NSW) and collected under the authority of the <i>Census and Statistics Act 1905</i> (Cth).	N	N	Y
NSW Emergency Department Data Collection	From 2006 - 2022	for all persons aged 25 years and over at the time of the 2016 census (9 August 2016) who have had a CVD event/procedure in any of the multiple diagnosis or procedure fields).	disclosed by NSW Ministry of Health to the ABS under the authority of section 27B of the <i>Privacy and Personal Information Protection Act 1998</i> (NSW) and under section 10(1)(f) and 11(1)(f) of Schedule 1 of the <i>Health Privacy Principles in the Health Records and Information Privacy Act 2002</i> (NSW) and collected under the authority of the <i>Census and Statistics Act 1905</i> (Cth).	N	Indigenous	Y

STATE AND TERRITORY DATA CUSTODIANS						
NSW Notifiable Conditions Information Management System	From 2020 - 2021	Anyone with a confirmed case of COVID-19 (regardless of age).	disclosed by NSW Ministry of Health to the ABS under the authority of section 27B of the <i>Privacy and Personal Information Protection Act 1998</i> (NSW) and under section 10(1)(f) and 11(1)(f) of Schedule 1 of the <i>Health Privacy Principles in the Health Records and Information Privacy Act 2002</i> (NSW) and collected under the authority of the <i>Census and Statistics Act 1905</i> (Cth).	N	N	Y
QLD Department of Education						
Once-off linkages						
QLD National Assessment Program – Literacy and Numeracy (NAPLAN) data	From 2010 to 2018	Outcomes from the NAPLAN tests for Queensland state school students who were in Year 3 during 2010, 2011 and 2012.	Education General Provisions Act 2006	N	Children	N

STATE AND TERRITORY DATA CUSTODIANS						
SA Department for Education						
Once-off linkages						
NDDA South Australian (SA) Linkage: SA National Assessment Program Literacy and Numeracy (NAPLAN) 2008-2019	From 2008 to 2019	Information on education for South Australians, including school attendance and attainment outcomes.	Section 13 of the South Australian Public Sector (Data Sharing) Act 2016 and the South Australian Intra-Government Data Sharing Agreement (Ref B911 484). Census and Statistics Act 190 (Cth)	N	Children	N
NDDA South Australian (SA) Linkage: SA School Enrolment Census	From 2005 to 2019	Information on education for South Australians, including school attendance and attainment outcomes.	Section 13 of the South Australian Public Sector (Data Sharing) Act 2016 and the South Australian Intra-Government Data Sharing Agreement (Ref B911 484). Census and Statistics Act 190 (Cth)	N	Children, Aboriginal and Torres Strait Islander	Y
SACE Board of South Australia						
Once-off						
NDDA South Australian (SA) Linkage: South Australian Certificate of Education (SACE)	From 2014 to 2019	Information on education for South Australians, including school attendance and attainment outcomes.	Section 15(1)(m) of the SACE Board of South Australia Act 1983. South Australian Intra-Government Data Sharing Agreement (Ref B911 484). Census and Statistics Act 190 (Cth).	N	Children, Disability, Aboriginal and Torres Strait Islander	Y

STATE AND TERRITORY DATA CUSTODIANS						
University Admissions Centre						
Once-off linkages						
Universities preferences and offers, ATAR Calculations	From 2003	Student university preferences based on ATAR calculations	NSW Research Exemption under section 27B of the Privacy and Personal Information Protection Act 1998 (NSW)	Y	Children	N
VIC Department of Health and Human Services						
Once-off linkages						
Transmission and Response Epidemiology (TREVl)	1 January 2020 – 30 June 2024	TREVl is a fully digital, fully implemented contact tracing system that contains public health information related to COVID-19 as part of the Victoria COVID tracing platform. The TREVl system covers the whole program of contact tracing and COVID-19 case management in Victoria, including: report of test and test results, outbreak notifications, application of public health orders and interviews.	Secondary use or disclosure of health information is pursuant to exemptions for contained in Victorian Health Privacy Principles (HPP) 2.2(c), (f) and (h).	Y	Aboriginal and Torres Strait Islander, Disability	Y
Victorian Cancer Registry	1982 - 2023	Data items include patient information, cancer type and incidence, diagnosis information including notifying hospital and mortality.		Y	Disability, Aged	Y

STATE AND TERRITORY DATA CUSTODIANS						
Victorian Linkage MAP	From 1991 to 2018	The Victorian Linkage Map (VLM) is a system of linked records that are identified as belonging to the same person across 30 different Victorian health and human services datasets. A selection of VLM datasets have been linked to MADIP, including: Alcohol and Drug Information System; Birth Registry; Child Protection and Out of Home Care; ChildFIRST; Death Index; Dental Health Program Dataset; Disability Services; Family Services; Family Violence Services; Homelessness Services; Mental Health Community Support Services; National Disability Insurance Scheme Dataset; Public Housing Applications; Public Housing Tenancies; Public Mental Health Services; Sexual Assault Services; Victorian Admitted Episodes Dataset; Victorian Emergency Minimum Dataset; Youth Justice Data	Health Records Act 2001 (Vic) & the Privacy & Data Protection Act 2014 (Vic)	N	Children, Aboriginal and Torres Strait Islander, Disability, Aged	Y

STATE AND TERRITORY DATA CUSTODIANS						
VIC Department of Treasury and Finance						
Once-off linkages						
Integrated DTF dataset	2011-2018	This dataset includes information and data on property sales, valuations and property planning permits.	A Letter of Exchange between the ABS and Department of Treasury and Finance, Victoria (DTF) to approve the transfer of data was signed in 2019.	Y	Aboriginal and Torres Strait Islander	N

PRIVATE SECTOR DATA CUSTODIANS						
Australia and New Zealand Intensive Care Society (ANZICS)						
Once-off linkages						
Adult Patient Database (APD)	2017-2024	The ANZICS APD contains deidentified data about individuals admitted to intensive care units, including demographics, health, characteristics of their admission, care and discharge, as well as characteristics of the relevant intensive care unit. The APD does not contain the names of individuals and will be integrated by using a statistical linkage key	The data custodian is sharing deidentified information which is no longer considered personal information under the Privacy Act 1988.	Y	Disability, Aboriginal and Torres Strait Islander	N
Electricity Distributors and Embedded Networks						
Once-off linkages						
Electricity Meter data	From 2016-2021	This data contains dwelling level Electricity Meter data, including solar installations and solar hot water systems data from Clean Energy Regulators.	Data Sharing agreement under <i>Clean Energy Regulator Act 2007</i>	Y		N

PRIVATE SECTOR DATA CUSTODIANS						
National Insurance Data Asset						
Once-off linkages						
Private sector insurance data	2023	The data asset will provide Australian Government to understand household insurance availability, coverage and affordability and the intersection with climate and natural hazard risks and projections to help decision makers target disaster risk reduction and climate adaptation activities. Insurance data will be collected at the address level and linked to dwelling.	Census and Statistics Act 1905	Y		N
SQM Research						
Once-off linkages						
Rental vacancies data	2016, 2021	Dwelling level data purchased from SQM research for use with Census enumeration	Census and Statistics Act 1905	N	N	N

PRIVATE SECTOR DATA CUSTODIANS						
National Centre for Vocational Education Research (NCVER)						
Once-off linkages?						
Total VET Activity		3 NCVER supersession tables (2015 – 2022) for TVA data These are csv files as follows: 1. Program supersession 2. Subject supersession 3. RTO supersession	Census and Statistics Act 1905	Y	N	N