



Maddocks

Lawyers
Level 1
40 Macquarie Street
Barton ACT 2600 Australia

Telephone 61 2 6120 4800
Facsimile 61 2 6230 1479

info@maddocks.com.au
www.maddocks.com.au



Australian Bureau of Statistics

2022 MADIP PIA Update

Privacy Impact Assessment Update

Date of analysis: 21.03.22

Date of report finalisation: 25.05.22

Contents

Part A	EXECUTIVE SUMMARY	3
1.	Introduction	3
2.	This PIA Update process	4
3.	Summary of findings	4
4.	Recommendations	6
Part B	METHODOLOGY AND ASSUMPTIONS	11
5.	Our methodology	11
6.	Assumptions and qualifications	12
Part C	Overview of MADIP	14
7.	Context.....	14
8.	Overview of 2022 MADIP PIA Update	16
Part D	DESCRIPTION OF CHANGES AND ANALYSIS OF RISKS	18
9.	New types or categories of data being included, or potentially considered for inclusion, in MADIP	18
10.	Increased volume of data	26
11.	Expanded outputs	29
12.	Changed or new data handling processes	31
Part E	GLOSSARY	36
Attachment 1	Material reviewed	38
Attachment 2	List of Stakeholders consulted	39
Attachment 3	Summary of APP Compliance Risk	41
Attachment 4	Datasets linked with MADIP	42
Attachment 5	MADIP Information Flow	61

Part A EXECUTIVE SUMMARY

1. Introduction

- 1.1 The Multi-Agency Data Integration Project (**MADIP**) is a secure, person-based, research data asset that combines broad sets of information about Australian citizens, to facilitate the use and re-use of public data for research purposes.
- 1.2 ABS has engaged Maddocks, on behalf of the MADIP Board, to conduct this PIA update to consolidate the findings from PIA update processes for MADIP since publication of the MADIP PIA Update report in November 2019, as well as consider several additional planned changes.
- 1.3 The following changes and functionalities are covered in this **2022 MADIP PIA Update**:
- 1.3.1 new types or categories of data being included, or potentially considered for inclusion, in MADIP, where they will be prepared for linkage in accordance with existing MADIP processes;¹
 - 1.3.2 increased volume of data (in terms of increased data from new datasets, increased variables, and increased frequency of update of information) being included, or proposed for inclusion, in MADIP;
 - 1.3.3 expanded outputs being permitted as part of approved MADIP research projects; and
 - 1.3.4 several new or changed data handling practices including:
 - (a) ABS internal operational use of MADIP data in statistical collections and use outside the DataLab;
 - (b) permitting international researchers to access MADIP data in the DataLab on a case-by-case basis;
 - (c) streamlining the data custodian approval model so that data custodians can permit the ABS to approve the use of data for selected projects on their behalf;
 - (d) development of a user portal for the DataLab; and
 - (e) permitting data custodians to use the mechanisms available under the *Data Availability and Transparency 2022* (Cth) Act (**DAT Act**), to provide datasets to the ABS for inclusion in MADIP.
- 1.4 This 2022 MADIP PIA Update report considers the privacy impacts of the proposed updates, as well as any planned changes, and whether these updates will affect ABS' continued compliance with the *Privacy Act 1988* (Cth) (**Privacy Act**) and in particular, the Australian Privacy Principles (**APPs**), in respect of MADIP.

¹ We note that a new 'type' of data can also be described as a new 'category' of data (these terms are used interchangeably in this PIA report). Whether new data introduced into MADIP constitutes a 'new type or category' of data is based on an assessment of the nature, scope, context, purpose and risk of the data.

2. This PIA Update process

- 2.1 Undertaking a PIA 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.
- 2.2 Since the original 2018 MADIP PIA process was completed, there have been several subsequent processes to consider the potential privacy impacts of specific projects or changes, which were subject to separately documented privacy impact threshold assessments or PIAs. These can be summarised diagrammatically as:

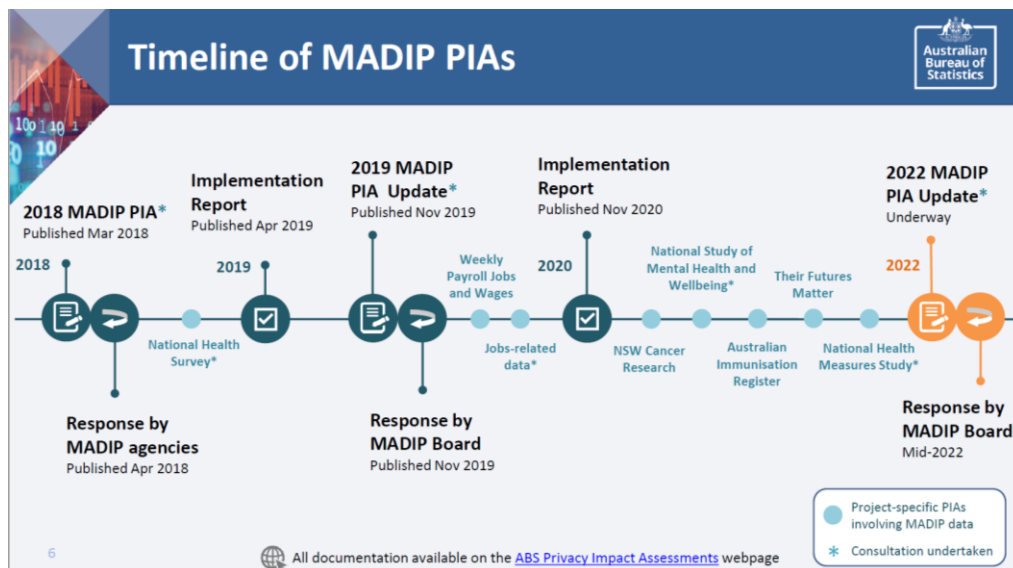


Image obtained from 2022 MADIP PIA Consultation Slides prepared by the ABS

- 2.3 Undertaking subsequent supplementary processes to re-examine the potential privacy impacts of an APP entity’s undertaking as plans and circumstances change, is entirely consistent with privacy best practice in Australia.
- 2.4 This 2022 MADIP PIA Update considers the privacy impacts of each of the changes and functionalities identified in paragraph 1.3 above, including:
 - 2.4.1 any effect on compliance with the Privacy Act, including the APPs;
 - 2.4.2 any changes to the previously identified information flows;
 - 2.4.3 any new privacy risks arising out of the changes, and any recommended mitigation strategies that could be implemented to address those risks.
- 2.5 This 2022 MADIP PIA Update is intended to help ABS and the MADIP Board manage any new or increased privacy risks and impacts, and we hope may also serve to inform stakeholders about the on-going importance of privacy considerations being a critical consideration for MADIP.

3. Summary of findings

- 3.1 In our view, the process of undertaking this 2022 MADIP PIA Update helps to demonstrate the ABS and MADIP Board’s commitment to thoroughly considering the privacy impacts of changes to the MADIP infrastructure and processes. We consider that it is consistent with privacy best practice in Australia to ensure that the general governance of MADIP includes

regular and comprehensive PIA update processes (such as this one) in addition to conducting privacy threshold assessments, and separate PIA processes for MADIP developments likely to have a significant impact on the privacy of individuals, as required by the APP Code.

3.2 In particular, we have been grateful for the ABS' role in co-ordinating and leading the extensive stakeholder consultation processes for this 2022 MADIP PIA Update. This has resulted in the extremely valuable contributions being received from stakeholders about the actual and perceived privacy impacts and risks of the new changes and functionalities that are within the scope of this PIA Update, and whether MADIP's existing mitigation strategies are appropriate and sufficient to address those risks. Again, we consider that conducting this broad stakeholder engagement process, which sought differing views to inform this PIA Update process, is consistent with privacy best practice.

3.3 While we have not identified any significant privacy risks that do not have existing mitigation strategies, we have identified:

3.3.1 that existing governance arrangements may not sufficiently facilitate proper consideration of risks or impacts of including new datasets into MADIP, including:

- (a) potential privacy and other risks for Aboriginal and Torres Strait Islander peoples² who are the subject of the relevant data;
- (b) the overall size of the MADIP asset, and whether adding the dataset would significantly increase the value of the asset for those who would seek to access and misuse the information in it (motivated intruder risk);
- (c) whether the existing process at the project proposal stage will be sufficient to address the risks that would, or could, arise if the new dataset was to be permitted to be linked (either on an enduring or 'one-off' basis) with other datasets within MADIP;

3.3.2 the need for additional mitigation strategies at the research project proposal and approval stages, including:

- (a) the need for further clarification about the intended purposes of MADIP, to assist data custodians in determining whether a specific research proposal falls within these purposes;
- (b) consideration about whether, and if so how, relevant stakeholders should be involved during the project proposal and approval process, including consumer groups and Aboriginal and Torres Strait Islander peoples;
- (c) ensuring that there are appropriate strategies in place, if it is proposed that any overseas researchers will have access to data in the DataLab for a specific research project; and
- (d) ensuring that, if any access outside the DataLab is proposed for a research project (such as in the ABS secure ICT environment), the data custodians can have confidence about the security of those arrangements;

3.3.3 the need to ensure that if access arrangements change, users of the DataLab portal continue to be clearly informed about how their personal information will be collected, used and disclosed; and

3.3.4 the need for continual (and potentially enhanced) transparency about MADIP, its datasets and approved projects.

² In this PIA Update, we have used the phrase 'Aboriginal and Torres Strait Islander' to describe individuals who identify as an Indigenous Australian, to respect the term that we understand most Aboriginal and Torres Strait Islander people prefer to use, but noting that in Australia there are many Indigenous nations, languages and cultures.

3.4 These risk and findings are considered in more detail throughout this PIA Update report. The recommendations set out in paragraph 4 of this **Part A** are designed to address the identified risks and further enhance privacy protections, and/or further strengthen compliance with the APPs.

4. Recommendations

4.1 We make the following recommendations in relation to the 2022 MADIP PIA Update:

<u>Recommendation 1</u> Transparency and implementation of recommendations from previous PIA processes	Relevant APPs
<p>Recommendation 1(a): We recommend that the MADIP Board consider whether it could take any additional steps to ensure that it has implemented the recommendations from previous PIA processes, when it has indicated that it will do so in its response to that recommendation.</p> <p>One way this could be done would be to prepare a consolidated implementation progress report (similar to that prepared by the ABS for some specific PIA processes), which documents:</p> <ul style="list-style-type: none"> • each recommendation that has been made; • the response to that recommendation at the time it was made (indicating whether there was an intention to implement that recommendation); and • whether (and if so how) the recommendation has been implemented or not yet implemented, or reasons why it is now not considered necessary to implement that recommendation. <p>As well as being a useful governance document (particularly if it is maintained going forward), and if published as a way of further enhancing transparency about MADIP, we suggest this document may be very useful in assisting the MADIP Board to identify any privacy trends or commonalities across the various PIA processes.</p> <p>Recommendation 1(b): As an additional transparency measure, we recommend that the MADIP Board take steps to ensure that there is transparency about the inclusion of new types or categories of data within MADIP, including:</p> <ul style="list-style-type: none"> • providing notice to the public that the MADIP Board is considering including new types or categories of data within MADIP, and calling for input from stakeholders about the impacts of any such inclusion (either through PIA processes or other consultations processes); and • following publication of the notice, publication of the reasons for including or not including new types or categories of data. 	<p>All</p>



<u>Recommendation 2</u> Criteria for including datasets in MADIP	Relevant APPs
<p>Recommendation 2(a): As an additional transparency and governance measure, we recommend that the MADIP Board consider developing and publishing criteria that will be used by the MADIP Board to assess whether a new dataset should be included within MADIP. This could be part of a broader framework or ‘guiding principles’ document, published on the MADIP website.</p> <p>At a minimum, we suggest that those criteria should include consideration of:</p> <ul style="list-style-type: none">• the public benefit of including the dataset within MADIP:<ul style="list-style-type: none">○ we believe this should be an overall assessment, taking into account all relevant factors including the other criteria listed below;○ if the public benefit test is not met, the dataset should not be included in MADIP;○ guidance should be prepared about how the ‘public benefit’ will be considered and assessed by the MADIP Board;• the utility of including the dataset within MADIP:<ul style="list-style-type: none">○ datasets should not be included in MADIP unless it is very likely that they will be both proposed by researchers for use in approved projects, <i>and</i> that those proposals will be approved in accordance with MADIP processes;• the types of information within the dataset and restrictions on use, including:<ul style="list-style-type: none">○ whether the dataset includes sensitive information (as defined in the Privacy Act), or other information likely to have a particular sensitivity (even if does not meet the Privacy Act definition);○ whether the dataset includes detailed information about businesses that may be sensitive if considered personal information;○ whether the dataset includes information about Aboriginal and Torres Strait Islander peoples or other vulnerable populations;○ whether the dataset is subject to particular legislative restrictions;○ whether data custodians have placed restrictions on use of the dataset;• the impact of including the dataset on the size and nature of the overall MADIP asset, including:<ul style="list-style-type: none">○ whether the inclusion of the dataset would increase the value of the asset for those who would seek to access and misuse the information. <p>We believe that these impacts could be assessed through a risk assessment process.</p> <p>Recommendation 2(b): We further recommend that relevant stakeholders (such as consumer groups and/or groups representing Aboriginal and Torres Strait Islander peoples) be consulted prior to finalising the criteria.</p>	<p>APP 1, APP 3, APP 6, APP 11</p>

<p><u>Recommendation 3</u> Including new types or categories of data in MADIP</p>	<p>Relevant APPs</p>
<p>In recognition of community perceptions that some types or categories of data are more sensitive than others, we recommend that:</p> <ul style="list-style-type: none"> • Recommendation 3(a): in addition to the criteria for including datasets in MADIP which would be developed under Recommendation 2, for any new type or category of data (for example, property insurance data), consider whether there should be additional governance or restrictions imposed for linking and use of that specific type or category of data within MADIP. For example: that a particular type or category of data would not be permitted to be linked to particular other datasets within MADIP; • Recommendation 3(b): in addition to the above, before including a new type or category of data, the MADIP Board should consider whether it is appropriate to seek input from relevant stakeholders, including from consumer groups. 	<p>APP 1, APP 3, APP 6</p>

<p><u>Recommendation 4</u> Data governance of data relating to Aboriginal and Torres Strait Islander peoples</p>	<p>Relevant APPs</p>
<p>Recommendation 4(a): We recommend that the MADIP Board work to ensure there is clear governance surrounding use of data about³ Aboriginal and Torres Strait Islander peoples, to provide certainty for Aboriginal and Torres Strait Islander people, but also clarity for data custodians and researchers seeking to use such data for research projects.</p> <p>We recommend that the MADIP Board continue to work with relevant stakeholder groups to develop that governance framework. We suggest that this is likely to need to explore a range of options, including:</p> <ul style="list-style-type: none"> • strategies for considering proposals for research projects that ensure that impacts on Aboriginal and Torres Strait Islander peoples are considered at this stage and that decisions are made in a properly informed way. This might involve: <ul style="list-style-type: none"> ○ specifying particular types of data or projects within MADIP, that require express consideration of Aboriginal and Torres Strait Islander issues; ○ requiring documentation for project proposals to include specific risks for Aboriginal and Torres Strait Islander peoples; ○ requiring such documentation to provide evidence that advice has been sought from appropriate persons or entities and/or the relevant Aboriginal and Torres Strait Islander communities that are the subject of the data have been consulted; • options for the MADIP Board to obtain relevant advice, for example: <ul style="list-style-type: none"> ○ from a bespoke or other advisory group; ○ from an ethical board; • options for representatives who identify as Aboriginal and Torres Strait Islanders being part of governance structures or arrangements; and/or 	<p>APP 1</p>

³ Use of data about Aboriginal and Torres Strait Islander peoples refers to information which is about and may affect Indigenous peoples both collectively or individually. See <https://www.maiamnayriwingara.org/key-principles>

<ul style="list-style-type: none"> consideration of frameworks in place overseas such as the <i>Ngā Tikanga Paihere Guidelines</i> prepared by Statistics New Zealand.⁴ <p>Recommendation 4(b): We appreciate that development and implementation of such a framework is likely to take time, if it is to be done in a truly consultative and genuinely engaged manner. We therefore recommend, as an interim option, the MADIP Board consider:</p> <ul style="list-style-type: none"> engaging an advisory body or individual to assist with the Aboriginal and Torres Strait Islander data issues; and working with that body or individual in respect of one or more specific proposed research projects that will involve the use of Aboriginal and Torres Strait Islander data, to co-design in consultation with key Aboriginal and Torres Strait Islander stakeholders and then trial models of engagement or other approval processes for that specific project. Learnings from such trials could then be used in developing the broader framework. 	
---	--

Recommendation 5 Additional strategies for research projects proposal and approval stages	Relevant APPs
<p>Recommendation 5(a): We recommend that the MADIP Board reconsider and reiterate the intended purposes of MADIP, to provide further clarity about ‘research purposes’ and whether this can include, for example, ABS operational use, and/or use to supplement surveys.</p> <p>Recommendation 5(b): We recommend that the MADIP Board ensure that, in addition to existing processes (including those that ensure the proposed linked datasets will be able to be de-identified before being included in the DataLab), the consideration and approval processes for research projects include:</p> <ul style="list-style-type: none"> consideration about whether, and if so how, relevant stakeholders should be involved in the project approval process, noting that depending on the relevant MADIP datasets, this may include involvement of consumer groups and Aboriginal and Torres Strait Islander people (as discussed in Recommendation 4); if any access to the linked datasets is proposed to occur outside of the DataLab (e.g. within the ABS secure IT environment so that additional tools that cannot be deployed in the DataLab can be used), a requirement for evidence that that environment is secure and that the linked datasets will be appropriately protected, and that access arrangements to the linked datasets are appropriately restricted; when implementing processes to address data quality, it particularly considers the data quality of the information in the MADIP datasets that are proposed to be linked for the project (particularly considering the difficulties in using datasets which contain data about Aboriginal and Torres Strait Islander peoples, given that there may be different reasons for, and types of, identification at different times and across datasets for Aboriginal and Torres Strait Islander peoples); if any access to the linked datasets is proposed to be given to individuals who will be physically located outside of Australia, ensuring that: 	<p style="text-align: center; font-weight: bold;">All</p>

⁴ The Ngā Tikanga Paihere Guidelines can be found at: <https://www.data.govt.nz/assets/data-ethics/Nga-Tikanga/Nga-Tikanga-Paihere-Guidelines-December-2020.pdf>.

<ul style="list-style-type: none"> ○ the risks associated with that access are formally considered and assessed as part of the project approval process (including reputational risks for MADIP if there is any inappropriate use or misuse of the data, particularly overseas where the protections of the Privacy Act and the Undertaking given for the purposes of the <i>Census and Statistics Act 1905</i> may not apply); and ○ appropriate additional mitigation strategies are put in place to address those risks, which might include (as required): <ul style="list-style-type: none"> ▪ additional contractual protections (noting the difficulties in enforcing international agreements); ▪ additional technical protections (such as increased monitoring of researchers' use within the DataLab, so that any inappropriate activity can be detected and prevented at an early stage); or ▪ additional security protections (such as restricted access arrangements or limiting tools available within the DataLab). 	
---	--

<u>Recommendation 6</u> Implementation of User Portal for the DataLab	Relevant APPs
<p>If access arrangements for the DataLab change (including the introduction of the user portal), we recommend that the MADIP Board take steps to ensure that appropriate collection notices continue to be reviewed and updated as necessary and continue to be displayed to users when they apply for access to, and use, the user portal, so that users are aware about how the ABS handles their personal information.</p>	<p>All</p>

Part B METHODOLOGY AND ASSUMPTIONS

5. Our methodology

- 5.1 This PIA Update has been conducted to identify and consider any new or enhanced privacy impacts and risks associated with the MADIP changes and new functionality within the scope of this PIA Update (that is, the matters outlined in paragraph 1.3 of **Part A [Executive Summary]** and in **Part D [Description of Changes and Analysis of Risks]**.
- 5.2 We conducted this PIA Update broadly in accordance with the *Privacy Impact Assessment Guide (PIA Guide)* issued by the Office of the Australian Information Commissioner (**OAIC**) in relation to undertaking privacy impact assessments. This involved the following steps:

Stage	Description of steps
1.	<p>Plan for the PIA: We have reviewed relevant background material provided by ABS, as outlined in Attachment 1, and were provided with briefings from ABS officers. We discussed the various changes that have been made, or are being considered to be made.</p> <p>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 the 2022 MADIP 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 ABS' proposed changes to MADIP. This draft was refined and then finalised following feedback from ABS.</p>
3.	<p>Stakeholder consultation: Undertaking consultation with stakeholders was considered an essential part of conducting this PIA Update. It provided both an opportunity to inform stakeholders about the updates and proposed changes to MADIP and the data integration infrastructure in ABS, and to listen to stakeholder views on the proposed updates and ABS' privacy management arrangements.</p> <p>We assisted ABS to prepare and plan the 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>Over 50 organisations participated in stakeholder consultation sessions for the 2022 PIA Update. We attended and participated in the consultation sessions, before preparing a report that summarises the consultations taken and the feedback provided to the ABS. We produced a Consultation Report, which summarises the feedback received. A copy of the Consultation Report will be published on ABS' website.</p>
4.	<p>Privacy impact analysis and compliance check: In this step we focussed on identifying the privacy impacts of each change, and its compliance against the relevant APPs and privacy best practice. In undertaking our analysis, we considered and applied the <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 compliance with the Privacy Act. 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, 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 compliance risk	<i>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.</i>
	Compliance could be further enhanced	<i>This rating indicates 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.</i>
	Improvements to meet best practice	<i>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.</i>
	Compliant	<i>This rating indicates that we have concluded that no further mitigation steps are required.</i>
A summary of overall compliance risk ratings against the APPs is set out at Attachment 3 .		
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 the draft PIA Update report: Following review of the draft report by ABS, we further refined our analysis and potential mitigation strategies as required to ensure that privacy risks were appropriately considered and addressed.	
8.	Report: We finalised this PIA report.	

- 5.3 The MADIP Board will review this PIA report and separately respond to our recommendations.
- 5.4 A glossary of defined terms and acronyms is at **Part E** of this PIA Update report.

6. Assumptions and qualifications

- 6.1 The scope of this PIA Update report is limited to the proposed or potential changes to MADIP as described in **Part D [Description of Changes and Analysis of Risks]**.
- 6.2 We have conducted our analysis on the basis that the factual information provided by ABS (as set out in **Part D [Description of Changes and Analysis of Risks]**) is up-to-date, correct and complete.
- 6.3 We have undertaken our analysis on the basis of the development of MADIP as at the date of analysis indicated in the cover page for this PIA Update report. We note and support the MADIP Board’s governance arrangements that will ensure that privacy impacts of future changes continue to be carefully considered, including through privacy threshold

assessments, and additional or supplementary PIA processes as required by the APP Code. In our view, this will:

- 6.3.1 enhance protections against the risk of “function creep”, which is an inherent privacy risk for projects which involve personal information being held and which endure over a significant period of time; and
- 6.3.2 provide a mechanism for considering the potential reforms that may arise out of the Australian Government’s current processes to review the Privacy Act (for example, changes that may result in new or more detailed requirements for information to be considered ‘de-identified’).

Part C Overview of MADIP

7. Context

- 7.1 This section seeks to set out, at a high level, an explanation about MADIP, its aims and objective, and describes in general terms some of the existing mitigation strategies that have already been deployed to ensure compliance with the Privacy Act. However, we have not sought to include all of the protections (which are discussed at length in reports resulting in the previous PIA processes), rather our aim is to give a general overview to help explain our further analysis in **Part D [Description of Changes and Analysis of Risks]**.
- 7.2 The Multi-Agency Data Integration Project (**MADIP**) is a secure, person-based research 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 data asset to facilitate the use and re-use of public data for research purposes. MADIP was initiated in 2015 to enhance, through use and re-use of public data, the evidence base for social policy.
- 7.3 MADIP is a cross-portfolio government partnership of seven Commonwealth agencies, being: Australian Bureau of Statistics (**ABS**), Australian Taxation Office (**ATO**), Department of Health (**Health**), Department of Education, Skills and Employment (**DESE**), Department of Home Affairs (**Home Affairs**), Department of Social Services (**DSS**), and Services Australia (**SA**).
- 7.4 Representatives from each of these agencies are members of the MADIP Board, which is responsible for the strategic direction of MADIP to ensure it continues to adapt to the changing data environment. MADIP is also supported by the MADIP Technical Advisory Group, which consults with analysts and custodians about technical arrangements, such as infrastructure, access arrangements and costing models.
- 7.5 The ABS is the accredited integrating authority for MADIP and is responsible for combining the datasets in MADIP, for providing access to MADIP data to only those who have been authorised (**Authorised Users**) for approved research projects, for ensuring the security of data contained within MADIP, and ensuring that all research results and other outputs from the use of MADIP data are not likely to identify any individual.
- 7.6 MADIP is bound by the constraints of:
- 7.6.1 the legislation of the data custodians that applies to data that they provide to MADIP;
 - 7.6.2 the *Census and Statistics Act 1905* (**Census and Statistics Act**);
 - 7.6.3 the Privacy Act; and
 - 7.6.4 social licence and community expectations.

Benefits of MADIP

- 7.7 MADIP is a secure research data asset used by authorised researchers for approved projects in a secure, virtual access environment. MADIP presents benefits to Australian agencies, the Australian public, researchers, and data providers (known as data custodians), including by:
- 7.7.1 providing Australian governments with a powerful tool for informing government decision-making;

- 7.7.2 supporting decisions that will help Australians live healthier, happier, and more independent lives;
- 7.7.3 making better use of the information that has already been collected to enhance the value of existing public data resources; and
- 7.7.4 making a wider range of data available for researchers from government, universities and public policy institutes.

MADIP data

- 7.8 MADIP contains high-value, person-centred and regularly updated datasets that aim to comprehensively cover the Australian population. The data contained within MADIP is currently primarily made up of Commonwealth data that is provided to the ABS for MADIP 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).
- 7.9 ABS has built a central linkage infrastructure within the MADIP. It does this by linking the datasets supplied by data custodians and entities to a central linkage infrastructure called the 'Person Linkage Spine' (**Spine**), whereby instead of linking one dataset to another, the Spine creates the capacity for separate source datasets to be linked where required for research projects. 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 projects. The Spine enables information to be brought together in relation to people who were residents in Australia during a given reference period, and currently contains information ranging from January 2006 to June 2021. The existing datasets contained in MADIP are depicted in the diagram at **Attachment 4**. An information flow diagram is set out in Figure 1 in **Attachment 5**.
- 7.10 MADIP adheres to the 'separation principle' and 'functional separation'⁵ in receiving, storing, and curating data for all integration projects. This means that for each dataset, data that contains identifying information about each individual that is the subject of the dataset ('linkage information'), is stored and handled separately from other data about that individual ('analytical information'). Linkage information is stored separately to de-identified, analytical information, and access to these different information sets is restricted so that no individual is able to access both sets of information simultaneously. In addition, functional separation means that ABS staff members undertaking data linkage only have access to the information they need to perform their assigned roles. Further information about how the separation principle is applied is set out in **Attachment 5**.
- 7.11 ABS is responsible for ensuring that the MADIP data that it makes available to authorised researchers in the secure DataLab is provided in a manner that is not likely to enable the identification of an individual (and therefore meets the requirements to be 'de-identified' under the Privacy Act)⁶.
- 7.12 ABS currently has a range of security arrangements in place for the IT systems used for the MADIP to protect MADIP data, which:
 - 7.12.1 conform with security arrangements set out in the Australian Government Information Security Manual (**ISM**);
 - 7.12.2 ensure that data collection, linkage, and assembly activities for MADIP datasets are only conducted by a dedicated team in the Secure Data Integration Environment (**SDIE**);

⁵ 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>.

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

- 7.12.3 includes a secured internet gateway which is reviewed annually by the Australian Signals Directorate (**ASD**); and
 - 7.12.4 includes an ongoing program of security audits and system accreditations, including the Information Security Registered Assessors Program (**IRAP**).
- 7.13 Further, MADIP datasets are handled in accordance with a range of additional privacy protection practices, including:
- 7.13.1 the Five Safes Framework, an internationally recognised approach to managing disclosure risks, which is applied to ensure access to MADIP 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;
 - 7.13.2 data may also be added to MADIP 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 the Spine;
 - 7.13.3 the ABS is transparent about linkages with MADIP and approved projects that make use of MADIP data through registers on the ABS website.

8. Overview of 2022 MADIP PIA Update

- 8.1 ABS regularly assesses the privacy risks of MADIP, and has undertaken previous privacy impact assessments (**PIA**) that have been published on the ABS website⁷, including:
- 8.1.1 original MADIP PIA report in March 2018 (**2018 MADIP PIA**);
 - 8.1.2 MADIP PIA Update report in November 2019 (**2019 MADIP PIA Update**);
 - 8.1.3 Cloud DataLab PIA report in June 2020;
 - 8.1.4 Jobs related data integration project PIA report in August 2020; and
 - 8.1.5 AIR-MADIP data integration project PIA report in August 2021.
- 8.2 While we support the conduct of these processes, and note that they are all published on the MADIP website, we note that it is somewhat difficult to track the recommendations that have been made in different processes, and ascertain whether or not the recommendations have been implemented. Accordingly, we **recommend** that the MADIP Board consider whether it could take any additional steps to ensure that it has implemented the recommendations from previous PIA processes, when it has indicated that it will do so in its response to that recommendation (**Recommendation 1**).
- 8.3 One way this could be done would be to prepare a consolidated implementation progress report (similar to that prepared by the ABS for some specific PIA processes), which documents:
- 8.3.1 each recommendation that has been made;
 - 8.3.2 the response to that recommendation at the time it was made (indicating whether there was an intention to implement that recommendation); and

⁷ <https://www.abs.gov.au/websitedbs/d3310114.nsf/home/abs+privacy+impact+assessments>

- 8.3.3 whether (and if so how) the recommendation has been implemented or not yet implemented, or reasons why it is now not considered necessary to implement that recommendation.
- 8.4 As well as being a useful governance document (particularly if it is maintained going forward), and if published a way of further enhancing transparency about MADIP, we suggest this document may be very useful in assisting the MADIP Board to identify any privacy trends or commonalities across the various PIA processes.
- 8.5 Each of the following changes and functionality are within the scope of this 2022 MADIP PIA Update as are described in more detail in **Part D [Description of Changes and Analysis of Risks]**:
- 8.5.1 new types or categories of data being included, or potentially considered for inclusion, in MADIP, where they will be prepared for linkage in accordance with existing MADIP processes;
- 8.5.2 increased volume of data being included, or proposed for inclusion, in MADIP;
- 8.5.3 expanded outputs being permitted as part of approved MADIP research projects; and
- 8.5.4 several new or changed data handling practices including:
- (a) ABS internal operational use of MADIP data in statistical collections and use outside the DataLab;
 - (b) permitting international researchers to access MADIP data in the DataLab on a case-by-case basis;
 - (c) streamlining the data custodian approval model so that data custodians can permit the ABS to provide approval on their behalf;
 - (d) development of a user portal for the DataLab; and
 - (e) permitting data custodians to use the mechanisms available under the DAT Act to provide datasets to the ABS for inclusion in MADIP.

Part D DESCRIPTION OF CHANGES AND ANALYSIS OF RISKS

9. New types or categories of data being included, or potentially considered for inclusion, in MADIP

Current state

9.1 The MADIP asset grows and evolves in response to Australia's information needs and changes in the broader data environment. To improve the quality and coverage of the Australian population, and therefore the value of MADIP, MADIP is updated regularly with more recent data, and the introduction of new datasets or the expanded linkage of some datasets. Responsibility for approving new data linkages to MADIP belongs to:

9.1.1 the MADIP Board and the relevant custodians of the data, where the data linkages are enduring; and

9.1.2 the ABS and the relevant custodians of the data, if the data linkage is once-off.

New changes and functionality

9.2 Since the 2019 MADIP PIA Update, the number of datasets linked to the Spine has expanded, and the ABS is proposing further expansions through the inclusion of additional datasets. Further, since the last PIA Update was undertaken, the ABS has proposed to expand some of the current MADIP data linkages, such as with the Business Longitudinal Analysis Data Environment (**BLADE**).

9.3 BLADE is a statistical resource that contains information on Australian businesses. It is a series of data modules and methodologies for linking business datasets by using a de-identified Australia Business Number (**ABN**) as the identifier. BLADE includes information about business tax data, government program data and information from ABS surveys.

9.4 Currently, only 'high-level business characteristics' from BLADE (known as BLADE core module) are linked with MADIP. For example, currently BLADE core module is linked to migration and Census data in MADIP to support analysis of the characteristics of businesses that benefit from employer sponsored migration. The Vaccines Strategy Project links the industry code from BLADE core module to the Australian Immunisation Register (**AIR**) dataset. It is used to analyse vaccine uptake rates by occupation and industry, especially within the aged care and disability sectors.

9.5 There has been an increased demand for analysis of detailed business level information with person level information to enable the same linkage to be analysed from different perspectives, that is, using the business or the person as the unit of analysis.

9.6 This proposed new type of data would enable BLADE business data with an increased level of detail to be linked with MADIP (rather than only high-level business characteristics as already linked with MADIP). Linking additional data sources from BLADE to MADIP will expand the areas of investigation that researchers can explore from an employer-employee perspective. This will allow the ABS to meet a growing area of demand, while still

maintaining the security of both the personal and business level data as described in this PIA.

9.7 The potential new and expanded data linkages include:

- 9.7.1 Further expansion of the link between MADIP and BLADE, to enable linkage from business to person and person to business, such as:
- (a) the creation of an analytical flag to note if a Registered Training Organisation (**RTO**) appearing within the Total VET Activity dataset is a sole trader. This flag would support privacy considerations in analysis of Total VET Activity data in MADIP, by allowing data specific to sole traders to be suppressed from output if necessary. This flag would be enabled by merging public data (i.e. RTO IDs and ABNs) from training.gov.au with BLADE core to ascertain the 'Type of Legal Organisation' status that can be assigned to an RTO appearing in the Total VET Activity dataset. This process is a demonstration of a consumer (student) and business (RTO) relationship within select BLADE and MADIP data, rather than the traditional employee to employer relationship; and
 - (b) further State and Territory government data, such as:
 - (i) the Industry and Employment Dynamics in South Australia project, which seeks to link South Australia government Business Support and Grants programs data to BLADE/MADIP;
- 9.7.2 Potential for linking new types or categories of data. For example:
- (a) property insurance data (to identify the levels of insurance coverage of the Australian population and improve the information base for managing natural hazards and climate risks); and
 - (b) expanded use of detailed Aboriginal and Torres Strait Islander specific data (depending on the research objectives).
- 9.7.3 We acknowledge that the MADIP Board does not have any specific proposals to link private sector categories of data to MADIP at this time. However, the concept was included as part of the stakeholder consultation for this 2022 MADIP PIA Update to start identifying what additional privacy steps would be necessary or desirable for this type of data if it were considered in the future. Consultation reinforced that, if the MADIP Board were to consider linking private sector categories of data to MADIP, the ABS will need to undertake further consultation and consideration of privacy issues to inform the consideration of the inclusion of the new type or category of data.

Analysis and recommendations

- 9.7.4 In **Table 1** below we identify privacy impacts and risks associated with the 'new types or categories of data', and our analysis against relevant APPs. We have also identified, where relevant, some of the existing mitigation strategies that address the risk or impact.

Table 1: New Types or categories of Data for Linkage – Analysis and Recommendations

<p>Impacts and Risks, including as identified during consultation (including assessment of compliance risk rating)</p>	<p>Existing mitigation strategies</p>	<p>Analysis and Recommendations</p>
<p>1.1 Individuals do not consent to the use of their information by MADIP</p> <p>[Improvements to meet best practice]</p>	<p>Statutory constraints on ABS on its role as accredited integrating authority for MADIP.</p> <p>MADIP existing governance framework requires that before a dataset is included in MADIP, it is assessed to ensure that specific Australian Privacy Principles (APP), in particular APP 5 and APP 6, are met.</p> <p>Data custodians are responsible for ensuring there is appropriate authorisation for the disclosure of relevant personal information to the ABS for MADIP. Data custodians may rely on an exception in APP 6.2 or APP 6.3 for the provision of data to ABS for MADIP, including APP 6.2(a), that the individual would reasonably expect the APP entity to use or disclose the information for the secondary purpose.</p>	<p>Stakeholders expressed a concern that inclusion of new datasets in MADIP were occurring without the specific consent of the individuals about whom the data relates, and questioned more generally how MADIP processes took the concerns of consumers into account.</p> <p>This issue relates to APP 6, which provides that an APP entity must not use or disclose personal information that was collected for a primary purpose for another purpose (a secondary purpose), unless the individual has consented to the use or disclosure of the information, or an exception in APP 6.2 or APP 6.3 applies.</p> <p>We note that APP 6.2(b) enables the use and disclosure of personal information by an APP entity if that use or disclosure of the information is required or authorised by or under an Australian law. In the context of MADIP, this generally means the law under which the data is collected that authorises disclosure (but it may include authorisations under the Census and Statistics Act). We note the constraints in the Census and Statistics Act, which apply to MADIP, including outcomes of data linkages as a result of MADIP.</p> <p>We further note, that if the ABS were to request consent from individuals for data linkages within MADIP, where consent to link data has been sought, but not given, this could lead to the risk of resulting data being biased (and therefore undermining the utility of MADIP). We note the existing mechanisms in the MADIP governance framework go some way to address this risk.</p> <p>However, we consider that further governance and transparency measures can be undertaken which would inform persons about how datasets are incorporated into MADIP, which would assist in addressing the underlying stakeholder concerns about persons not being informed about how their data is used for data linkage. We have set out these measures in Recommendation 1 and Recommendation 2.</p>

<p>Impacts and Risks, including as identified during consultation (including assessment of compliance risk rating)</p>	<p>Existing mitigation strategies</p>	<p>Analysis and Recommendations</p>
<p>1.2 The linking of private sector datasets to MADIP may have a greater impact on individuals and introduce risks not currently contemplated</p> <p>[Improvements to meet best practice]</p>	<p>No specific measure for private sector data, but existing governance framework as per above. This includes conducting a PIA to consider and reduce any privacy risks of including new types or categories of data in MADIP.</p> <p>Existing governance framework for consideration of research projects and deidentified outcomes from research projects.</p>	<p>We note that the MADIP Board does not have any specific proposals to link private sector categories of data to MADIP at this time. However, the concept was included in the stakeholder consultation for this PIA to start identifying what additional privacy steps would be necessary or desirable for this type of data if it were considered in the future.</p> <p>We note that some stakeholders raised concerns about the potential inclusion of private sector data in MADIP. Stakeholders considered private sector data to be inherently different to government administrative data. In our view it is difficult to fully contemplate the risks that may flow from the linkage of private sector data without considering in detail a proposed dataset. We note that some private sector datasets are particularly sensitive, such as health insurance related data. In addition to considering whether personal information ‘can’ be collected and then used under the APPs (e.g. ABS ‘can’ collect the dataset because the collection meets the legislative test of being reasonably necessary, or directly related to one or more of ABS’s functions), we consider that it is also important that proper consideration be given to whether information ‘should’ be collected and then used (e.g. whether it is appropriate and safe to include the datasets within MADIP).</p> <p>We therefore recommend additional measures at Recommendation 3 (in addition to the measures in Recommendation 2) to guide consideration of the inclusion of new types or categories of data for example, property insurance data.</p>
<p>1.3 There is the potential that the inclusion of new datasets into MADIP may lead to greater risk of re-identification of individuals</p>	<p>ABS manages disclosure risk through implementation of the Five Safes Framework, which is used to support safe and effective access to microdata for Authorised Users, in the ABS DataLab.</p>	<p>We note that stakeholders were generally concerned about the potential for re-identification of individuals caused by increased numbers of datasets within MADIP.</p> <p>However, it is important to draw a distinction between the inclusion of new datasets within MADIP, and the provision of that data to researchers for approved projects. We suggest that the re-identification risk identified by stakeholders properly relates to the latter of these. We note the existing mitigation processes (including implementation of the separation principle and functional separation) designed to mitigate the risk of any re-identification during the linkage process.</p>



Impacts and Risks, including as identified during consultation (including assessment of compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
[Compliant]	<p>As part of the Five Safes assessment, the ABS Disclosure Review Committee considers the risk of re-identification and recommends steps to minimise disclosure risk.</p> <p>Researchers accessing the data are liable for criminal penalties if they misuse the data, such as seeking to identify individuals.</p> <p>Further, the Census and Statistics Act, which governs the collection, use and disclosure of data for research and statistical purposes, requires that the ABS cannot release information of a personal or domestic nature in a manner that is likely to enable the identification of that person.</p>	<p>We note that de-identification is a process which involves the removal or alteration of personal identifiers, followed by the application of any additional techniques or controls required to remove, obscure, aggregate, alter and/or protect data in some way so that it is no longer about an identifiable (or reasonably identifiable) individual. The potential for disclosure of personal information engages APP 6 and APP 11. In addition to being required for legal compliance (i.e. Privacy Act), using appropriate de-identification techniques can promote trust and meet community expectations around data handling practices, and alleviate public concerns about the handling of personal information. OAIC Deidentification Guidance provides the test for determining if an individual is 'reasonably identifiable' from otherwise de-identified data. Data is considered to be 'deidentified' where the risk of reidentification is very low.</p> <p>Some stakeholders were concerned with the re-identification risks that could arise if private sector data was linked with MADIP, especially if access arrangements were changed to allow private sector bodies access to MADIP data. However, it was confirmed that linking private sector data is not currently under consideration nor is any change to access arrangements to allow private organisations access to MADIP data. In our view, such changes would warrant a further examination of the privacy implications.</p> <p>We consider that the existing mitigation strategies, including application of the Five Safes Framework, are likely to be sufficient to mitigate against the risks of re-identification.</p>
1.4 There is a lack of self-governance for Aboriginal and Torres Strait Islander data, leading to data about Aboriginal and Torres Strait Islander	Not Applicable	<p>We note that the stakeholder consultations identified the handling of data about Aboriginal and Torres Strait Islander people as a significant issue. The consultations highlighted that it is particularly important to consider data sovereignty principles and risks in respect of data about Aboriginal and Torres Strait Islander peoples.</p> <p>We note that the use and disclosure of Aboriginal and Torres Strait Islander data may involve a risk that the Aboriginal and Torres Strait Islander people may not have a meaningful understanding of how their data will be handled via MADIP, and feel a loss of control in respect of the use and public</p>



<p>Impacts and Risks, including as identified during consultation (including assessment of compliance risk rating)</p>	<p>Existing mitigation strategies</p>	<p>Analysis and Recommendations</p>
<p>peoples potentially not being used in a sufficiently ethical and culturally-safe manner</p> <p>[Improvements to meet best practice]</p>		<p>release of information which is about them. A lack of proper consultation or clear governance about the including, and subsequent use, of Aboriginal and Torres Strait Islander data through MADIP leads to a risk of Aboriginal and Torres Strait Islander data not being used in an ethical or culturally-safe manner.</p> <p>Stakeholders in a specific Aboriginal and Torres Strait Islander consultation workshop strongly expressed the view that the Aboriginal and Torres Strait Islander population needs greater involvement to ensure proper control over their data, and emphasised the need for proper community consultation around using that data.</p> <p>Currently, the governance arrangements for MADIP have no express requirements that result in the opportunity for Aboriginal and Torres Strait Islander peoples to have input into how their data may be used (and/or outcomes from projects released). We note that the decision to include datasets within MADIP, or to approve for use for research projects, will be made by data custodians (or potentially by ABS on behalf of a data custodian, if the relevant data custodian has previously approved this). Under the current processes, there is limited opportunity for Aboriginal and Torres Strait Islander peoples to have input into or gain understanding of the ways in which Aboriginal and Torres Strait Islander data can be included in, and then used as part of, MADIP.</p> <p>We acknowledge that the issues raised at the consultation sessions are broader than considerations of ‘personal information’ as currently defined under the Privacy Act, and therefore beyond the scope of what can be fully addressed under this PIA Update process, which analyses compliance against the Privacy Act. We also acknowledge that issues about data sovereignty and ethical use of Aboriginal and Torres Strait Islander peoples’ data is a complex issue, and as the consultations for this PIA Update highlighted, there is unlikely to be a ‘one size fits all’ approach (as demonstrated by the range of options that were put forward at the consultation sessions by stakeholders).</p> <p>However, we note that APP 1 is intended to ensure that entities manage personal information in an open and transparent way. We therefore recommend in that context that the MADIP Board</p>



Impacts and Risks, including as identified during consultation (including assessment of compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
		consider a range of options to address the potential privacy impacts for Aboriginal and Torres Strait Islanders, as set out in Recommendation 2 , Recommendation 4 and Recommendation 5 .
<p>1.5 Aboriginal and Torres Strait Islander privacy may not be sufficiently protected at the community level</p> <p>[Improvements to meet best practice]</p>	Not Applicable	<p>Stakeholders noted that privacy and other impacts at a community level is a concern, and that this is especially true of projects considering data for remote areas, where specification of a particular location could be the equivalent of identifying a specific Aboriginal and Torres Strait Islander community. As such, individual privacy should be considered alongside community well-being for projects involving Aboriginal and Torres Strait Islander data. Relevant stakeholders (including Cultural Authorities) may assist with understanding privacy risks specific to each community.</p> <p>The nature of MADIP data is such that it contains sensitive information, including about individuals who are receiving social security benefits and/or participating in employment services programs. The use, and in particular the disclosure, of information about specific Aboriginal and Torres Strait Islander communities is likely to have an impact for those communities, and for those living as part of those communities.</p> <p>Again, we note that the Privacy Act focuses on protections for individuals, and does not acknowledge or protect ‘communities’. The OAIC does recommend that all entities consider their ethical and social obligations as part of the broader management of projects, even where the materialisation of these risks would not necessarily result in a breach of the Privacy Act.</p> <p>In this context, it is important to recognise the risk of ‘attribute disclosure’. This is where it is possible to learn something about a data subject, without actually identifying them. For example, the outcomes from a specific research project could incorrectly lead to assumptions about characteristics of a group, leading to harm such as reputational damage, deficit narratives or embarrassment for that group, even though it was not possible to identify any individual from the data.</p>



Impacts and Risks, including as identified during consultation (including assessment of compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
		<p>Attribute disclosure may not, in and of itself, constitute re-identification, however there may nevertheless be a risk of harm to members of a group due to Aboriginal and Torres Strait Islander community norms that emphasise kinship and interconnectedness.</p> <p>Accordingly, we set out a range of measures in Recommendation 2, Recommendation 4 and Recommendation 5 to address the concerns raised.</p>
<p>1.6 Increasing the number of datasets may increase the likelihood that motivated intruders will endeavour to inappropriate access and/or misuse the data</p> <p>[Compliant]</p>	<p>[Considered in more detail in Table 2 below]</p> <p>MADIP data is not stored altogether as a single dataset. It is only brought together for particular research projects as required.</p>	<p>This risk is considered in more detail in Table 2 below, and we have concluded that existing mitigations are likely to be sufficient to protect against this risk.</p> <p>However, we have suggested that this is a risk that should be considered by the MADIP Board when considering whether or not a new dataset should be included in MADIP (see Recommendation 2).</p>

10. Increased volume of data

Current state

- 10.1 As MADIP continues to expand, the volume of data (in terms of increased data from new datasets, increased variables, and increased frequency of update of information) in MADIP has increased with linkage of new datasets, and updated years of existing datasets.
- 10.2 There is also an increased demand for linked data that can be analysed at more frequent intervals than previously possible. There has been increasing demand for data from MADIP to be supplied at close to real time intervals to inform government policy priorities and meet researcher requirements. This requires the ABS to receive and share data at intervals of up to one week. For example, data supporting analysis for the AIR-MADIP project is made available to researchers within 2 weeks from the time of collection. This purpose of this project is to inform the national vaccine rollout by analysing vaccination rates and instances.
- 10.3 The ABS applies data minimisation as a core principle of data collection, use and access arrangements for MADIP. Only information that is reasonably necessary for projects is shared with ABS and linked as part of MADIP.

New state

- 10.4 To respond to evolving policy needs, like has been the case to inform the policy response to the COVID-19 pandemic, it is anticipated that in the future further datasets which include more recent data of individuals will be linked more frequently to MADIP.

Analysis and recommendations

- 10.5 In **Table 2** below we identify the privacy impacts and risks associated with the 'increased volume of data', and our analysis against relevant APPs. We have also identified, where relevant, some of the existing mitigation strategies that address the risk or impact.

Table 2: increased Volume of Data - Analysis and Recommendations

Impacts and Risks identified during consultation (including compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
<p>2.1 That the increased volume of data within MADIP will increase the likelihood of 'motivated intruders'</p> <p>[Compliant]</p>	<p>There are a range of security measures currently in place to protect MADIP, described at Part C sections 7.11 and 7.13.</p> <p>MADIP data is not stored altogether. It is only brought together for particular research projects as required.</p>	<p>APP 11.1 requires an APP entity to take such steps as are reasonable to protect personal information from misuse, interference and loss, and from unauthorised access, modification or disclosure.</p> <p>We note that as the volume of data increases within MADIP, there is greater incentive for 'motivated intruders' to attempt to break through the security measures protecting MADIP data to access personal information.</p> <p>We note that ABS has already undertaken a range of steps to meet its APP 11 obligations, including undertaking regular IRAP assessments and conforming with the ISM security requirements. We further note that MADIP data is handled through privacy protection practices including the Separation Principle and the Five Safes Framework.</p> <p>Given the above, we consider that ABS has taken reasonable steps to protect data within MADIP, including the proposed increased volume of data, and is therefore compliant with APP 11.</p>
<p>2.2 That the increased volume of data within MADIP will increase the likelihood that the data can be re-identified</p> <p>[Improvements to meet best practice]</p>	<p>There are existing measures for researchers accessing the MADIP data including:</p> <ul style="list-style-type: none"> • training for staff to look out for these specific issues; • auditing of DataLab sessions to ensure that best practice is being adhered to; and • undertakings signed by researchers which are legally binding documents specifying 	<p>We consider that as the number of datasets included within MADIP grows, there is greater opportunity to create and release different combinations of dataset linkages. As different combinations of datasets are available to Authorised Users, there is a heightened risk that the Authorised Users could use the different combinations of datasets to reidentify the data.</p> <p>Indeed, stakeholders in the consultation sessions expressed concern that the increased volume of data would increase risks that the data is identifiable. Noting the existing mitigation strategies, stakeholders queried whether there would be any changes made to the current disclosure and vetting arrangements to deal with re-identification risks.</p> <p>We consider that the implementation of Recommendation 2 and Recommendation 3 will enable the ABS to control the growth in volume of data within MADIP, and that existing processes at the research project consideration stage are sufficient to ensure that re-identification risks are addressed before any disclosure of MADIP data.</p>



Impacts and Risks identified during consultation (including compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
	conditions for access to the for access to the data. Penalties for misuse of data in the DataLab are up to \$25,200.	

11. Expanded outputs

Current state

- 11.1 MADIP facilitates the combination of linked datasets for researchers (Authorised Users) to analyse microdata, and to create aggregated outputs. To produce outputs, researchers are provided access to de-identified MADIP data through a data analysis solution that enables them to undertake analysis of microdata – referred to as the DataLab.
- 11.2 Only aggregated information is permitted to leave DataLab for broader use in accordance with the approved project objectives, and ABS staff check all data outputs to ensure that individuals cannot be reasonably identified through the aggregate outputs created, such as by unusual combinations of data.
- 11.3 The ABS uses the Five Safes Framework⁸ to manage access of information by authorised researchers. The Five Safes Framework requires the following questions to be answered prior to the disclosure of information:
- 11.3.1 Is the researcher authorised to access and use the data appropriately (safe people)?
 - 11.3.2 Is the data to be used for an appropriate purpose (safe projects)?
 - 11.3.3 Does the access environment prevent unauthorised use (safe settings)?
 - 11.3.4 Has appropriate and sufficient protection been applied to the data (safe data)?
 - 11.3.5 Are the statistical results non-disclosive (safe output)?

New state

- 11.4 There has been an increased demand for the use of MADIP data by ABS officers to produce official statistics. The ABS is also considering expanding the use of MADIP to inform ABS survey collections. There is also increased interest in having an aggregated, business level view of characteristics of persons associated with a business/organisation.
- 11.5 ABS is considering expanding MADIP outputs in the following ways:
- 11.5.1 by linking integrated MADIP data with ABS survey data collections in order to replace data items, such as healthcare, tax or childcare costs with data that is already available in MADIP⁹. This is to ease the burden on survey respondents and provide accurate information to enhance survey data and maximise the use of existing information held by government agencies; and
 - 11.5.2 by releasing aggregate personal information of consumers associated with an identified business, where such information is relevant to measuring business performance. For example, releasing post-training employment outcomes specific to training providers. This information is only available to ABS officers, not general MADIP researchers.¹⁰

⁸ Further details on the Five Safes Framework can be found at: <https://www.abs.gov.au/about/data-services/data-confidentiality-guide/five-safes-framework>.

⁹ This is currently likely to be done by transferring the relevant MADIP data from the Secure Data Integration Environment to a secure ABS processing environment for updating the survey record. See information flow diagram (**Attachment 45**). In some cases, the survey data may be recombined with MADIP for use as part of the MADIP integrated data asset and available to researchers.

¹⁰ Note: Information about training providers that are sole traders is not released to ensure the personal information of training providers is not disclosed.

Analysis and recommendations

- 11.6 The privacy impacts and risks associated with ‘expanded outputs’, and our analysis against relevant APPs, have already been described in **Table 1** and **Table 2** above.
- 11.7 In addition, we note that the protections ABS has in place mean that access is not provided to sole trader information, which ensures that information about a person is not disclosed in a way that could identify a person. There are protections in the Census and Statistics Act that ensure identified business information is not made available to researchers accessing MADIP data.
- 11.8 We consider that the mitigation strategies recommended in those Tables (that is, **Recommendation 2**, **Recommendation 3**, **Recommendation 4**, and **Recommendation 5**) will also work to address the identified impacts and risks.

12. Changed or new data handling processes

Current state

- 12.1 ABS and the MADIP Board continuously review and improve data handling practices and infrastructure for MADIP in order to preserve privacy, ensure data security, increase data quality and timeliness, and adhere to the governing legislation in relation to the data it holds.

New state

- 12.2 The ABS is considering potential changes and updates to the MADIP operating environment, including:
- 12.2.1 **change in MADIP information flow to facilitate ABS internal operational use of MADIP data in statistical collections**, by transferring MADIP data from the Secure Data Integration Environment (**SDIE**) and the ABS DataLab, to storing and using the data in the secure ABS IT environment to facilitate approved projects. Refer to diagram at **Attachment 4**);
 - 12.2.2 **permitting international researchers to access MADIP data in the DataLab on a case-by-case basis**, to enable international comparison studies or research that is of benefit to Australia. International researchers, who include experts in their field, would be subject to the same requirements as other approved researchers in the DataLab;
 - 12.2.3 **potential changes to the data custodian approval model to streamline the approval process and reduce burden on data custodians**. Through the addition of a threshold approval model (as negotiated with data custodians) the ABS would approve the use of the data for selected projects on behalf of the data custodian. In this approval model, criteria is applied to identify the types of projects that can make use of the data custodian's data without further consideration and the projects that will require delegate review and approval;
 - 12.2.4 **new user portal for the DataLab**: change to data flows and use of cloud storage for securely storing and linking data projects in DataLab were considered in the Cloud DataLab PIA. Further enhancements including development of a user portal to support MADIP will align with the PIA considerations;
 - 12.2.5 **exploring the possibility of introducing systems to enable the sharing of information to the ABS for MADIP through the mechanisms available under the DAT Act**.

Analysis and recommendations

- 12.3 In **Table 3** below we identify the impacts and risks associated with 'expanded outputs' and 'changed or new handling data practices' against the relevant APPs. We have also identified, where relevant, existing mitigation strategies that address the risk or impacts.

Table 3: Changed and new data handling processes - Analysis and Recommendations

<p>Impacts and Risks identified during consultation (including compliance risk rating)</p>	<p>Existing mitigation strategies</p>	<p>Analysis and Recommendations</p>
<p>3.1 That expanded types of research projects may not be wholly consistent with the MADIP framework</p> <p>[Improvements to meet best practice]</p>	<p>Current processes for considering whether Survey Data should be supplemented by MADIP data.</p>	<p>In our view, this is relevant to consideration of APP 1 (openness and transparency of data handling processes) as it is important that there is clarity for data custodians, for researchers, and for the public about MADIP, its purposes and permitted research projects.</p> <p>In stakeholder consultations, some data custodians identified the need for further clarity and guidance about the use of MADIP to support or facilitate operational use by an agency, and whether this was consistent with the original intentions for MADIP (i.e. use for research purposes). We recommend that the position be clarified by the MADIP Board as set out in Recommendation 5, to enhance transparency and openness in relation to MADIP and its datasets.</p>
<p>3.2 That overseas researchers gaining access to the DataLab may not comply with contractual obligations for APP compliance, resulting in misuse of MADIP data and adverse impacts for individuals</p> <p>[Compliance can be further enhanced]</p>	<p>Reputation of research entity (and individual researchers) is considered during project approval process.</p> <p>International researchers are subject to contract requiring compliance with the APPs, which could be legally enforced (even if this may be difficult in practice).</p> <p>Application of s 16C of the Privacy Act (which provides that, in certain circumstances, an act done or practice engaged in, by the overseas recipient is taken to be done or engaged in by the relevant APP entity).</p> <p>Additional conditions are considered on a case-by-case basis and include:</p>	<p>APP 8 is relevant to the disclosure of MADIP data to any individual located outside of Australia, even if the data remains on servers located within Australia (as is the case for the DataLab). It requires that, before any such disclosure occurs, such steps as are reasonable in the circumstances to be taken to ensure that the overseas recipient does not breach the APPs (unless an exception in APP 8.2 applies).</p> <p>While there are several mitigation strategies in place that will be such steps, we note that it may be difficult in practice to legally enforce contractual obligations with an overseas entity or individual, which stakeholders identified as a concern that such a contract may not act as an effective deterrent to prevent breaches of the APPs.</p> <p>Accordingly, we have recommended (in Recommendation 5) that the project approval processes for projects that involve access by overseas entities or individuals, include assessment of the size and nature of the risk of breach of an APP by that entity or individual (taking into account the nature of the entity/individual, and their knowledge and experience in Australian privacy issues) and the impacts for the individuals subject to the data if a breach were to occur (taking into account the data to which they will have access, and the risk that it could be used to reidentify and adversely impact individuals).</p>



Impacts and Risks identified during consultation (including compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
	<ul style="list-style-type: none"> tailored additional training specific to the international researcher; sponsorship of the research by an Australian agency or organisation; additional auditing of DataLab sessions. <p>International researchers will also be bound by any legal frameworks applicable in their respective nations, and any restrictions on the use of MADIP data in the context of conducting research from those countries would be subject to those laws. We consider that those restrictions would also be an effective deterrent for researchers from misusing MADIP data.</p>	<p>The MADIP Board could then ensure that appropriate additional strategies are put in place as required. Depending on the circumstances, these could include:</p> <ul style="list-style-type: none"> additional contractual protections; additional technical protections to what is usually undertaken (such as increased monitoring of researchers' use within the DataLab, so that any inappropriate activity can be detected and prevented at an early stage); or additional security protections (such as restricted access arrangements or limiting tools available within the DataLab).
<p>3.3 That ABS may not appropriately grant (or refuse) approval on behalf of a data custodian</p> <p>[Compliant]</p>	<p>Streamlined arrangements will only apply in accordance with documented criteria (as agreed with the relevant data custodian(s)) that sets the limits of the approval that can be granted by the ABS on behalf of the data custodian.</p>	<p>In our view, this measure does not raise significant privacy risks or impacts that will not be appropriately addressed by existing governance arrangements and proposed limitations of the streamlined approval process.</p>

Impacts and Risks identified during consultation (including compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
	<p>All other approvals require data custodian consideration and decision.</p> <p>Data custodians will be responsible for ensuring that the streamlined arrangements are appropriate for safe handling of data.</p>	
<p>3.4 New user portal [Improvements to meet best practice]</p>	<p>Only limited personal information about users will be collected, and only used to facilitate access to the DataLab for MADIP projects.</p>	<p>We understand that only the name and email address of the user will be collected, in order to create and store a user profile, that will be used to permit access to the DataLab. Although this does not appear to raise any significant privacy risks, we have not undertaken a full investigation and analysis of all aspects of how this information will be collected, stored, used and disclosed. We think it will be critical that users are properly advised about this, before they decide to provide information (for example, whether the collected information will be used to monitor not only access to the DataLab itself, but access to and use of data within the DataLab). We understand that users of the DataLab are already provided with a clear notice that explains how their personal information will be collected, and then used or disclosed, in connection with the DataLab. Accordingly, we have recommended that should arrangements change, suitable collection notices should continue to be displayed to users when they apply for access to, and available when they use, the User Portal (Recommendation 6).</p>
<p>3.5 Use of the DAT Act. [Compliant]</p>	<p>It is not proposed that DAT Act authorisation would be used to provide authorisation for ABS to share data <i>from</i> MADIP.</p> <p>The DAT Scheme is regulated by the National Data Commissioner.</p>	<p>Some stakeholders expressed concerns about the potential for the DAT Act to be used to authorise the sharing/disclosure of data from MADIP, but it was confirmed that this was not within the scope of current consideration (and that all existing processes would continue to apply). It was also clarified that data custodians may use the DAT Act authorisation to provide datasets to MADIP, but this would not be required (i.e. data custodians would have a choice about whether to use DAT Act authorisation or rely on other authorisation).</p>



Impacts and Risks identified during consultation (including compliance risk rating)	Existing mitigation strategies	Analysis and Recommendations
		Given this clarification (and the protections and requirements that we understand are included in the DAT Act before authorisation can be used), this measure does not raise significant privacy risks or impacts that will not be appropriately addressed by existing governance arrangements and mitigation strategies.

Part E GLOSSARY

Definitions	
ABS	means the Australian Bureau of Statistics.
ATO	means the Australian Taxation Office.
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> .
ASD	means the Australian Signals Directorate.
Authorised Users	means those who are authorised to access MADIP data.
BLADE	means the Business Longitudinal Analysis Data Environment that is proposed to be linked to MADIP.
Census and Statistics Act	means the <i>Census and Statistics Act 1905</i> (Cth).
DAT Act	means the <i>Data Availability and Transparency Act 2022</i> .
DataLab	means the secure analytical cloud environment that provides safe access for approved researchers to analyse de-identified microdata.
DESE	means the Department of Education, Skills and Employment.
DSS	means the Department of Social Services
Health	means the Department of Health.
Home Affairs	means the Department of Home Affairs.
IRAP	means the Information Security Registered Assessors Program, one of the security arrangements in place for the IT systems used for the MADIP.
ISM	means the Australian Government Information Security Manual, one of the security arrangements in place for the IT systems used for the MADIP.
OAIC	means the Office of the Australian Information Commissioner.
MADIP	means the Multi-Agency Data Integration Project, a secure, person-based research 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 re-use of public data for research purposes.
personal information	has the meaning given in section 6 of the Privacy Act.

Definitions	
PIA	means this privacy impact assessment.
PIA Guide	means the <i>Privacy Impact Assessment Guide</i> issued by the OAIC.
Privacy Act	means the <i>Privacy Act 1988</i> (Cth).
SA	means Services Australia.
SDIE	means the Secure Data Integration Environment.
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 MADIP that creates the capacity for separate source datasets to be linked.
2018 MADIP PIA	means the original MADIP PIA undertaken in 2018.
2019 MADIP PIA Update	means the MADIP PIA Update report in 2019.

Attachment 1 Material reviewed

1. ABS Access Process - Microdata, *received 18 November 2021*
2. MADIP five safes operational overview, *received 18 November 2021*
3. ABS DataLab integrated microdata project proposal, *received 18 November 2021*
4. MADIP update consultation plan, *received 18 November 2021*
5. MADIP legal framework, *received 18 November 2021*
6. MADIP governance arrangements, *received 18 November 2021*
7. MADIP strategy, *received 18 November 2021*
8. MADIP Data Availability and Access User Information Pack, *received 18 November 2021*
9. MADIP strategy on a page, *received 18 November 2021*
10. Initial project description – MADIP PIA Update [for Maddocks], *received 18 November 2021*
11. ABS Building Trust Strategy & MADIP Operating Model (in draft), *received 18 November 2021*
12. Datasets in MADIP, *received 22 November 2021*
13. OECD DataLab users approval, *received 22 November 2021*

Attachment 2 List of Stakeholders consulted

The list below sets out the organisations that attended the consultation sessions for the 2022 MADIP PIA Update:

1. ARC Centre of Excellence for Children and Families over the Life Course (The Life Course Centre)
2. Australian Bureau of Agricultural and Resource Economics and Sciences
3. Australian Climate Service
4. Australian Institute of Aboriginal and Torres Strait Islander Studies
5. Australian Institute of Health and Welfare
6. Australian National University
7. Australian Taxation Office
8. Barang Regional Alliance
9. Binarri Binyja Yarrowoo Empowered community
10. Cancer Council Australia
11. Cancer Institute of New South Wales
12. Centre for Big Data Research in Health, University of New South Wales
13. Commonwealth Scientific and Industrial Research Organisation / Data61
14. Consumers Health Forum of Australia
15. Department of Education, Skills and Employment
16. Department of Health
17. Department of Prime Minister and Cabinet
18. Department of Social Services
19. Empowered Communities (independent)
20. Grattan Institute
21. Healing Foundation
22. Information and Privacy Commission New South Wales
23. Kaiela Institute
24. Kimberley Aboriginal Medical Services
25. Maiaim Nayri Wingara Indigenous Data Sovereignty Collective
26. Menzies School of Health Research
27. Monash University
28. National Health and Medical Research Council
29. National Indigenous Australians Agency
30. Notitia Consulting (aligned to Empowered Communities)
31. Ngaanyatjarra Pitjantjatjara Yankunytjatjara Empowered Community
32. Office of the Australian Information Commissioner
33. Office of the Information Commissioner Northern Territory
34. Office of the Victorian Information Commissioner
35. Pama Futures (Far North Queensland Empowered Communities)
36. Privacy Committee of South Australia
37. Queensland Government Statistician's Office
38. Queensland Department of Health
39. Services Australia
40. South Australian Chief Data Officer
41. South Australian Department of Premier and Cabinet
42. Universities Australia
43. University of Sydney
44. Victorian Aboriginal Community Controlled Health Organisation
45. Victorian Centre for Data Insights
46. Victorian Department of Jobs, Precincts and Regions
47. Victorian Department of Premier and Cabinet
48. Western Australia Department of Communities

49. Western Australia Department of Premier and Cabinet Aboriginal Engagement Unit
50. Winnunga Nimmityjah Aboriginal Health Service

Attachment 3 Summary of APP Compliance Risk

The table below summaries the APP compliance risk rating for the changes and functionalities covered by the 2022 MADIP PIA Update based on the descriptors set out in **Part B**.

Australian Privacy Principle (APP)	Changes and Functionalities covered in the 2022 MADIP PIA Update			
	New types or categories of data for linkage	Increased volume ¹¹ of data	Expanded Outputs	Data handling practices
APP 1 - open and transparent management of personal information	Improvements to meet best practice	Improvements to meet best practice	Improvements to meet best practice	Improvements to meet best practice
APP 2 - anonymity and pseudonymity	Compliant	Compliant	Compliant	Compliant
APP 3 - collection of solicited personal information]	Compliance could be further enhanced	Improvements to meet best practice	Compliant	Compliant
APP 4 - dealing with unsolicited personal information	Compliant	Compliant	Compliant	Compliant
APP 5 - notification of the collection of personal information	Compliant	Compliant	Compliant	Improvements to meet best practice
APP 6 - use or disclosure of personal information	Improvements to meet best practice	Compliant	Compliance could be further enhanced	Compliant
APP 7 - direct marketing	Not applicable	Not applicable	Not applicable	Not applicable
APP 8 - cross-border disclosure of personal information	Not applicable	Not applicable	Not applicable	Improvements to meet best practice
APP 9 - adoption, use or disclosure of government related identifiers	Not applicable	Not applicable	Not applicable	Not applicable
APP 10 - quality of personal information	Improvements to meet best practice	Improvements to meet best practice	Compliant	Compliant
APP 11 - security of personal information	Compliant	Compliant	Compliant	Improvements to meet best practice
APP 12 - access to personal information	Compliant	Compliant	Compliant	Compliant
APP 13 - correction of personal information	Compliant	Compliant	Compliant	Compliant

¹¹ In terms of increased data from new data sets, increased variables, and increased frequency of update of information.

Attachment 4 Datasets linked with MADIP

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

Datasets currently linked to MADIP

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	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	Y		
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
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
Business Longitudinal Analysis	From 2001	BLADE is an economic data tool combining tax, trade, and intellectual	Census and Statistics Act 1905	Y		N

¹² 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 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
Data Environment (BLADE) – CORE Dataset: Indicative data items; BLADE Locations data; Business Activity Statement; Business Income Tax; Payment Summaries Pay As You Go		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 MADIP.				
Census of Population and Housing (Census)	2011, 2016	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
Once-off linkages						
ABS Household Income & Expenditure Survey	2015-16	Two phase survey of Survey of Income & House 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 Survey of Income and Housing	2019-20, 2020-21	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	Y		N
Australian Children’s Education and Care Quality Authority (ACECQA)						

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
Once-off linkages						
National Quality Standard	2013 to 2017	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 s 271 (1c)	Y		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).	Y	Children, Disability, Aboriginal and Torres Strait Islander	Y
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)	Y		N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
			enables the ATO to provide the ABS with unit record PIT data.			
Personal Income Tax (PIT) data: ATO Client Register, PAYG Payment Summaries, Client Data Income Tax Returns	From 2006 (Client Register), From 2010-11 (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
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).	Y		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	Y		N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
			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.			
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.	Y		N
Superannuation Balances data – Member Account Transaction Service (MATS) and Member Account Attributes Service (MAAS)	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.	Y		N
Taxable Payment Annual Report	2020-21	Total payments made to contractors by residential aged care facilities,	ATO data is disclosed to the Australian Statistician under the	Y		N

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
(TPAR) from Residential Aged Care Facilities, Payee Data		including subcontractors, consultants and independent contractors operating as companies, partnerships, trusts or sole traders.	Taxation Administration Act 1953 (Cth) for the 'purpose of administering the Census and Statistics Act 1905' under which the MADIP is conducted.			
Department of Defence and Department of Veterans' Affairs						
Once-off linkages						
Veterans Population Spine and Analytical Dataset	2001 to 2018	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). 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.	Y		Y
Department of Education, Skills and Employment						
Enduring linkages						
Australian Apprenticeships Incentives Program (AAIP) & Training Contracts	2006 to 2019	Information about trainees & apprentices, qualifications, their employers/trainers & incentive payments provided through the program	The AAIP is governed by the Australian Apprenticeships Incentives Program Guidelines, a document which is publicly available. Section I.E of the	N	Aboriginal and Torres Strait Islander, Disability	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
			AAIP Guidelines states the Apprenticeship Network Providers (who are contracted by the Department to deliver support services) 'may collect, disclose, make a record or otherwise use personal information for the purposes of administering the Program. The Privacy Act 1988 and the Australian Privacy Principles govern how personal is collected, used, disclosed, and stored.			
Australian Early Development Census (AEDC)	2009, 2012, 2015, 2018	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
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

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
Provider Registration and International Students Management System (PRISMS) (from 2006)	From 2006	Information on the course enrolments of overseas students.	Census and Statistics Act 1905	Y		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).	Y	Children, Aboriginal and Torres Strait Islander, Disability	Y
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 accordance 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-	2015-16 to 2018-19	Demographic and employment outcomes information collected from Job Seekers during the delivery of employment services.	Social Security (Administration) Act 1999, subsection 202(2C)	Y	Aboriginal and Torres Strait Islander, Disability	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
Program Monitoring Surveys (PPM)						
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	Y	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.	Y	Children	Y
Student Residential Address and Other Information (Student) collection	2018, 2019, 2020, 2021	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.	Y	Children	N
Department of Health						
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).	Y	Children, Aboriginal and Torres Strait Islander, Aged	Y
Centralised Register of Medical	From 2011	Information about registered medical practitioners, including specialities	Data is disclosed to the ABS pursuant to public interest certificates (PICs) issued by the	N	Children, Aged, Disability	Y

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
Practitioners (Provider Directory)			Minister (delegate) in accordance with section 130 of the Health Insurance Act 1972 and section 135A of the National Health Act 1953			
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
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

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
Once-off linkage						
Aged Care Provider data		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)	Y	Aged	N
Department of Home Affairs						
Enduring linkages						
Migration data – Adult Migrant English Program data	From 2003	Information on participants in the English language tuition program, including their visa type, course attendance and educational outcomes.	Census and Statistics Act 1905	Y		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
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						

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
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
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: s 208 of the Social Security (Administration) Act 1999; s 168 of the A New Tax System (Family Assistance) (Administration) Act 1999; s 128 of the Paid Parental Leave Act 2010; and s 355 of the Student Assistance Act 1973.	N	Aboriginal and Torres Strait Islander, Disability, Aged	Y
Once-off linkages						
Transgenerational Data Set (TDS) extract	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.	s 202(2C) of the Social Security (Administration) Act 1999; s 168(1)(a) of the A New Tax System (Family Assistance) Administration Act 1999; s 355(1)(a) of the Student Assistance Act 1973; & s 128 (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

COMMONWEALTH DATA CUSTODIANS						
DATASET NAME	REFERENCE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION ¹²	SENSITIVE DATA
National Disability Insurance Agency						
Enduring linkages						
National Disability Insurance Scheme (NDIS) Participant and Provider Data	From 2019	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).	Y	Children, Disability	Y
Services Australia						
Enduring linkages						
Medicare Consumer Directory	2006 to June 2020	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 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.	N	Children, Aged	Y

STATE AND TERRITORY DATA CUSTODIANS						
DATASET NAME	REFEREN CE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP 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	2006 to 2018	Information collected by State and Territory Registrars to enable the registration of a birth.	Statistics (Arrangements with States) Act 1956 (5)(1, 2)	Y	Aboriginal and Torres Strait Islander	Y
Cancer Institute of New South Wales						
Once-off linkages						
NSW Linked Cancer dataset: Cancer Incidence and Mortality	1972 to 2016	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	N	Aboriginal and Torres Strait Islander, Aged	Y

STATE AND TERRITORY DATA CUSTODIANS

DATASET NAME	REFEREN CE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION	SENSITIVE DATA
			Health that does not identify any individual to whom the information relates for statistical & research purposes.			
NSW Linked Cancer dataset: BreastScreen NSW	1988 to 2016	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: NSW Pap Test Registry	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
NSW Department of Communities and Justice						
Once-off linkages						

STATE AND TERRITORY DATA CUSTODIANS

DATASET NAME	REFEREN CE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION	SENSITIVE DATA
New South Wales Human Services Dataset	1990 to 2020	The HSIDS 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 HSIDS 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.	Y	Children, Aboriginal and Torres Strait Islander	Y
NSW Department of Education						
Once-off linkages						
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)	Y	Children	Y
New South Wales 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)	Y	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)	Y	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)	Y	Children	

STATE AND TERRITORY DATA CUSTODIANS						
DATASET NAME	REFEREN CE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION	SENSITIVE DATA
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)	Y	Children	
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)	Y		
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 2016	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)	Y	Children	
QLD Department of Education						
Once-off linkages						
QLD National Assessment Program – Literacy and Numeracy (NAPLAN) data	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
SA Department for Education						
Once-off linkages						
NDDA South Australian (SA) Linkage: SA National Assessment Program	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	Y	Children	N

STATE AND TERRITORY DATA CUSTODIANS						
DATASET NAME	REFEREN CE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION	SENSITIVE DATA
Literacy and Numeracy (NAPLAN) 2008-2019			the South Australian Intra-Government Data Sharing Agreement (Ref B911 484). Census and Statistics Act 190 (Cth)			
NDDA South Australian (SA) Linkage: SA School Enrolment Census	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)	Y	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)	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).	Y	Children, Disability, Aboriginal and Torres Strait Islander	Y
VIC Department of Health and Human Services						
Once-off linkages						
Victorian Linkage MAP	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	Health Records Act 2001 (Vic) & the Privacy & Data Protection Act 2014 (Vic)	N	Children, Aboriginal and Torres Strait Islander, Disability, Aged	Y

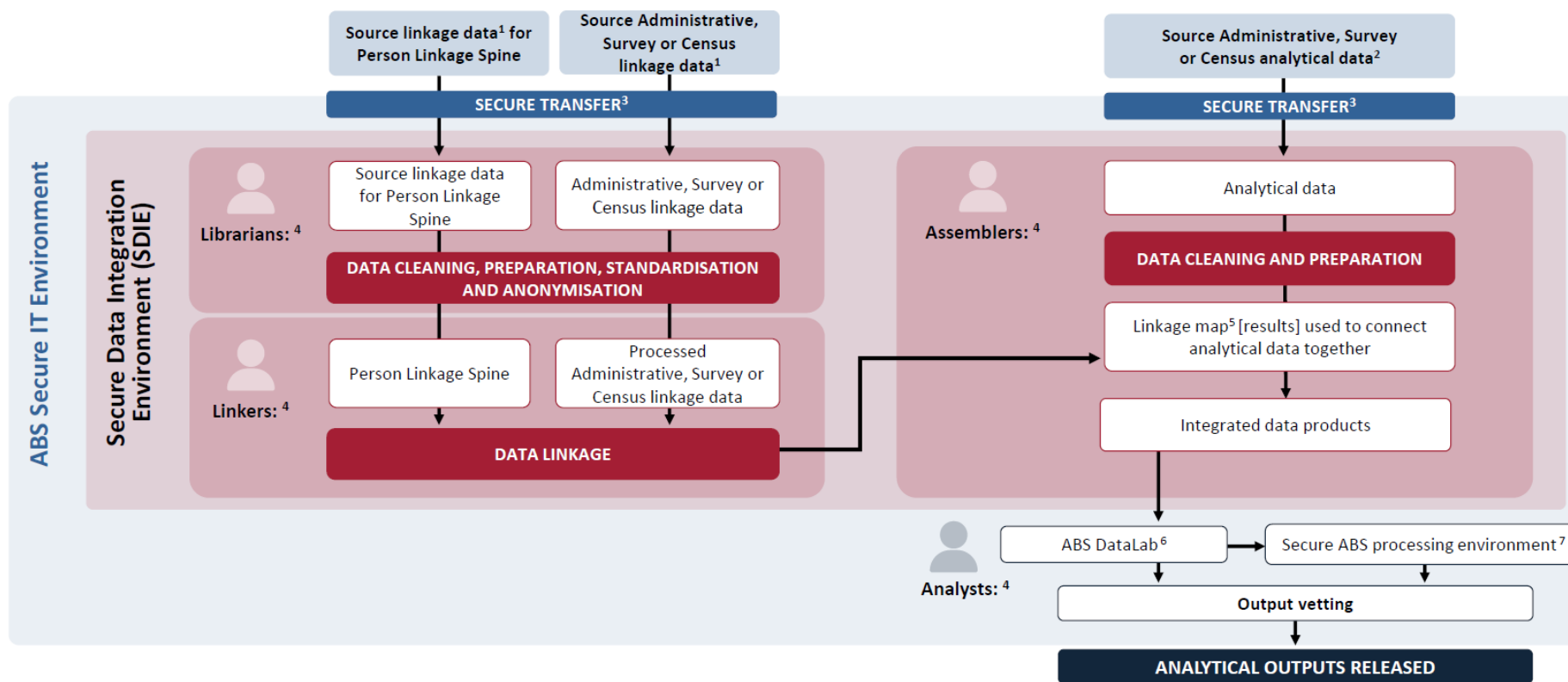
STATE AND TERRITORY DATA CUSTODIANS

DATASET NAME	REFEREN CE PERIOD	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	NEW SINCE 2019 MADIP PIA UPDATE	VULNERABLE POPULATION	SENSITIVE DATA
		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				

Attachment 5 MADIP Information Flow

[The following figures have been produced by ABS]

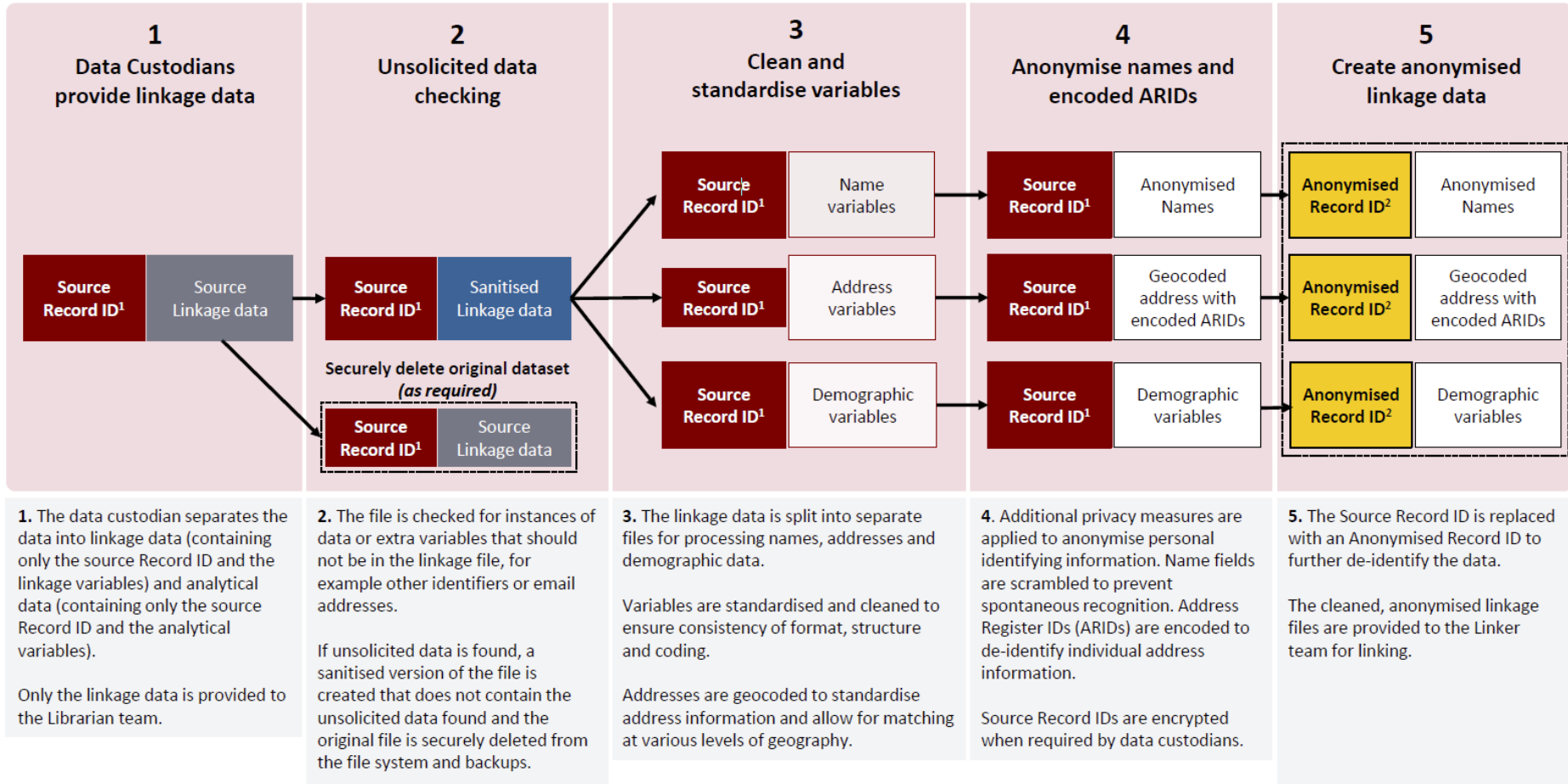
FIGURE 1: MADIP Data Flow



1. **Linkage data:** usually includes personal identifiers such as name, address and date of birth, or other identifiers like Australian Business Numbers. Note, some data items may be approved for both linkage and analytical use.
2. **Analytical Data:** variables of interest for analysis, such as occupation, income and health services use, or business type and industry.
3. **Secure transfer:** data is supplied to the ABS via an accredited secure transfer method as agreed with data custodians.
4. **Functional roles (librarians, linkers, assemblers and analysts):** The ABS applies the separation principle at all stages of the data process, ensuring that linkage data and analytical data are kept separate at all times. ABS officers are assigned to a specific functional role and are only able to access information necessary to perform their role. Refer to Figures 2, 3 and 4 for more information on the different roles undertaken by ABS staff.
5. **Linkage map:** Maps the links between IDs across datasets. IDs are generated by the ABS. Linkage maps contain no personal information. Refer to Figure 2.
6. **ABS DataLab:** a secure analytical cloud environment for users to undertake complex analysis of detailed microdata for statistical research or modelling. The ABS uses a series of controls as part of the Five Safes Framework to mitigate the risk of disclosure in the DataLab.
7. **Secure ABS processing environment:** a location within the secure ABS IT environment where MADIP data can be used by the ABS for approved projects. Access by ABS officers only in this location.

FIGURE 2: Librarian processes

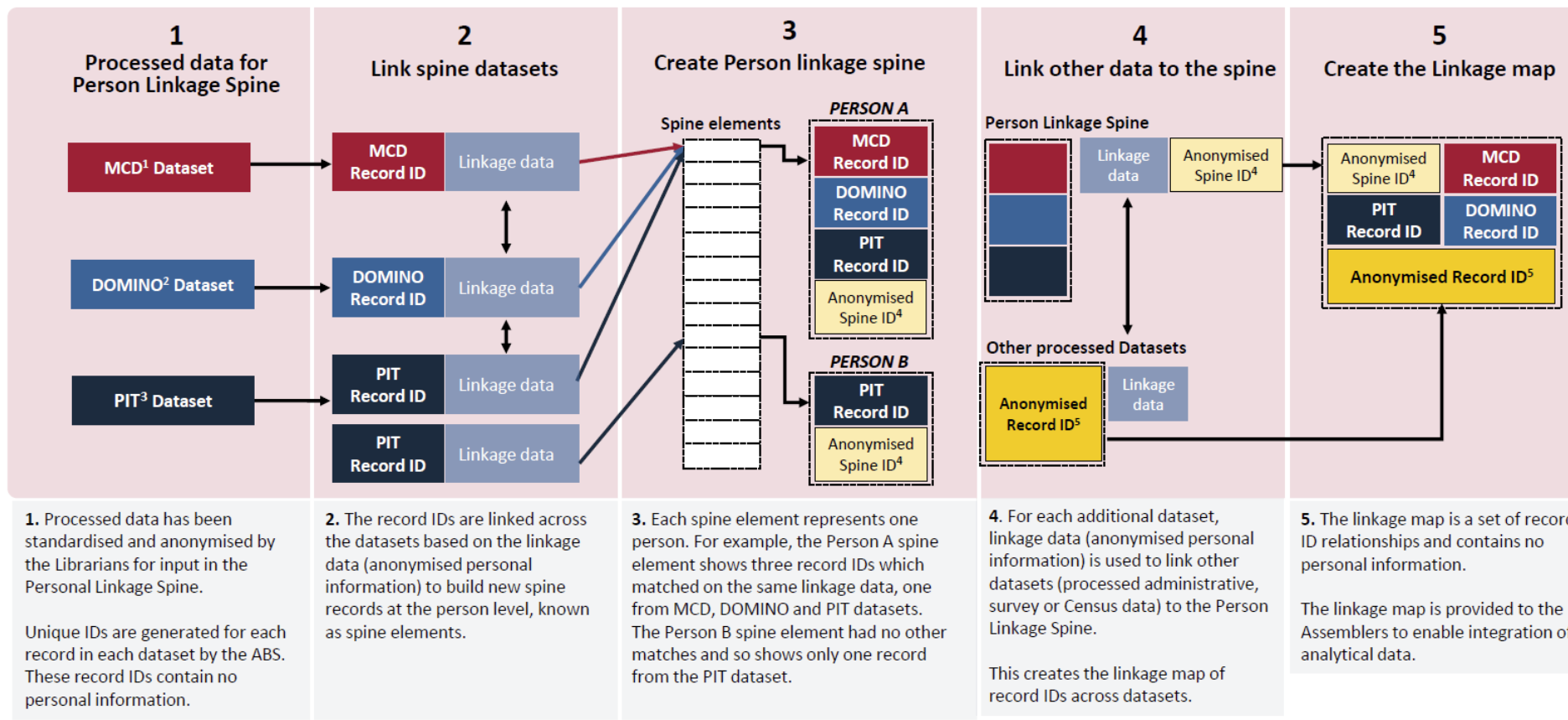
The diagram and process descriptions below outline the Librarian processes to create the anonymised linkage data.



1. Source Record IDs are sometimes anonymised by data custodians before they are provided to the ABS.
 2. Anonymised Record IDs are provided in the Linkage map.

FIGURE 3: Linker processes

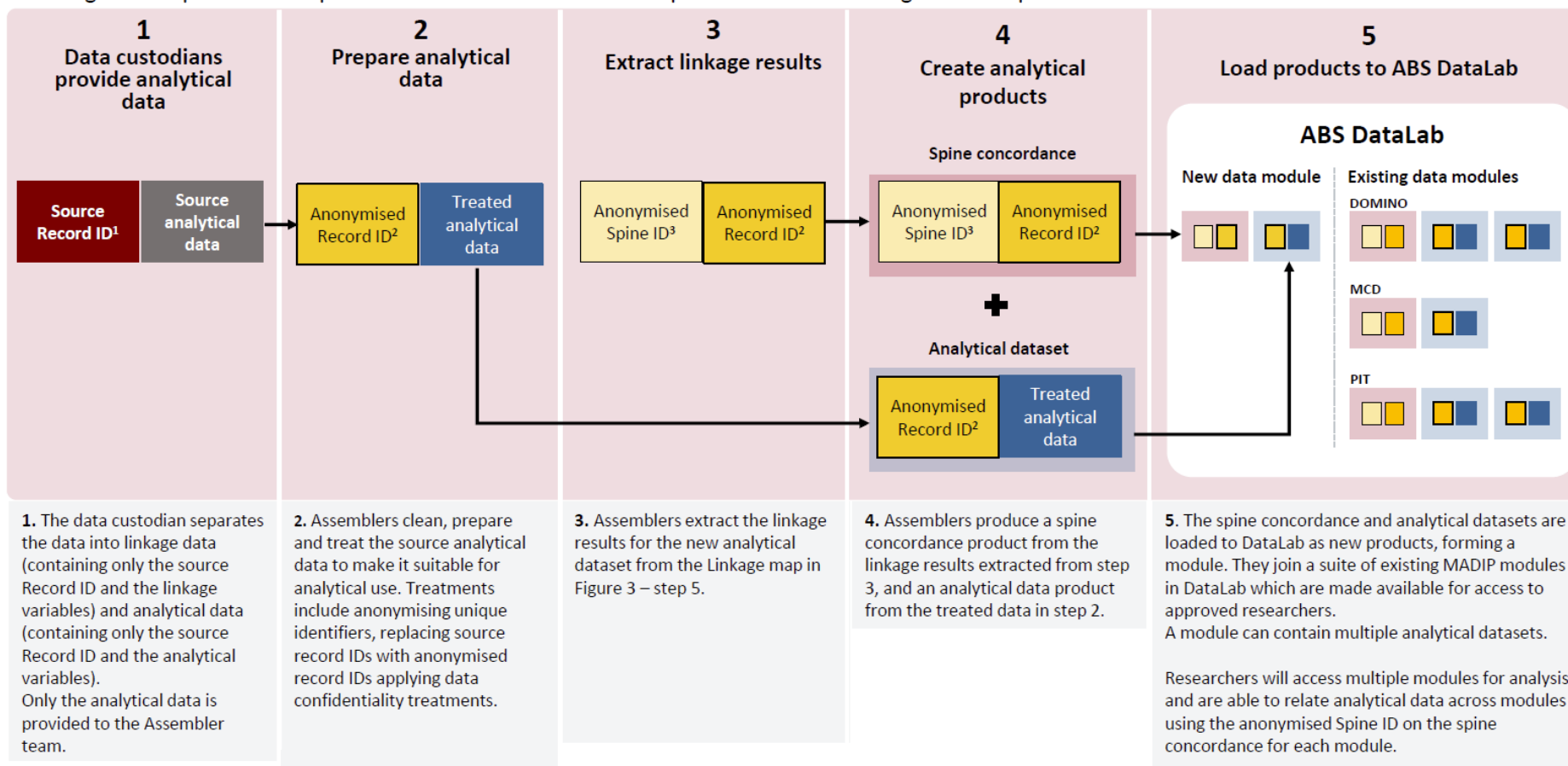
The Person Linkage Spine is the foundational linking infrastructure underpinning the MADIP data asset. The diagram and process description below outline the Linker processes to create the Person Linkage Spine and the Linkage Map.



1. Medicare Consumer Directory (MCD) dataset is provided by Services Australia on behalf of the Department of Health.
 2. Domino Centrelink Administrative Data (DOMINO) is provided by the Department of Social Services.
 3. Personal Income Tax (PIT) data is provided by the Australian Taxation Office.
 4. Anonymised Spine ID generated by the ABS contains no personal information
 5. Anonymised Record IDs are provided in the Linkage map and contain no personal information.

FIGURE 4: Assembler processes

The diagram and process descriptions below outline the Assembler processes to create integrated data products for access in the ABS DataLab.



1. Source Record IDs are sometimes anonymised by data custodians before they are provided to the ABS.
2. Anonymised Record IDs are provided in the Linkage map (see Figure 3)
3. Anonymised Spine ID generated by the ABS contains no personal information.

4. Medicare Consumer Directory (MCD) dataset is provided by Services Australia on behalf of the Department of Health
5. Domino Centrelink Administrative Data (DOMINO) is provided by the Department of Social Services
6. Personal Income Tax (PIT) data is provided by the Australian Taxation Office.