Australian Government
Australian Bureau of Statistics
Australian Taxation Office
Department of Education
Department of Health
Department of Human Services
Department of Social Services

MULTI-AGENCY **DATA INTEGRATION** PROJECT

Privacy Impact Assessment (PIA) Update for the Multi-Agency Data Integration Project (MADIP)

November 2019

CONTENTS

PART A – EXECUTIVE SUMMARY

1.1 Background

The Multi-Agency Data Integration Project (MADIP) commenced in June 2015 and became fully operational in June 2018. Research from MADIP provides whole-of-life insights that can improve the lives of all Australians. By bringing together a broad set of person-centred data from across government domains, MADIP facilitates the use and re-use of public data for research purposes. MADIP is a partnership among Australian Government agencies and operates under a framework of legislation, governance, and information management protocols that ensures data are shared and used for public benefit, privacy is protected, and data remains secure. The Australian Bureau of Statistics (ABS) is the accredited Integrating Authority for MADIP and, in collaboration with its partners, is responsible for combining the data, providing access to authorised users via highly secure ABS systems, and safeguarding privacy—ensuring that no individual person is likely to be identified. Data custodians, or entities authorised by data custodians, share data with the ABS for MADIP. The ABS does not release data from MADIP in a manner that is likely to identify any individual.

A Privacy Impact Assessment (PIA) for MADIP was independently undertaken by Galexia in late 2017-early 2018 and the report and response from MADIP Board members were published in April 2018. The report found MADIP to be broadly compliant with the Australian Privacy Principles (APPs) and made recommendations to improve governance, transparency, and data minimisation, storage, and retention. The MADIP Agencies' response agreed to the recommendations and committed to their implementation. An Implementation Report of the 2018 MADIP independent PIA was published in April 2019 to report on actions taken against each of the independent PIA recommendations.

1.2 Updating the MADIP PIA

MADIP is an evolving project. There have been a number of developments in MADIP since the previous PIA (the '2018 PIA') was published in April 2018. The key developments are:

- The amount of datasets linked through the MADIP environment has expanded, with additional years and new sources of data of the same broad type having been collected, integrated and stored using the MADIP framework.
- ABS survey data has been collected, integrated and stored (which was undertaken after a specific PIA was conducted using the MADIP framework).
- The ABS is using a new data system/environment for securely receiving, storing and linking MADIP data.
- The ABS has enhanced its person-centred data linkage infrastructure to more efficiently combine datasets on a project by project basis and strengthen privacy protections.
- The process for providing authorised researchers with access to MADIP data for research projects has been reformed and streamlined to deliver more timely access in a secure way.
- Consistent with the Australian Government's Secure Cloud Strategy, the ABS is planning to use secure cloud storage and computing for the DataLab.
- More detailed information about MADIP is available on the ABS website.

This PIA Update identifies, and provides useful assistance in managing the privacy risks associated with MADIP, including these developments since the 2018 PIA was conducted.

MADIP will continue to adapt and evolve to policy and research priorities, methodological and technological advancements, and other environmental changes. Some upcoming changes to MADIP include:

- The integration of further ABS survey data.
- A research project involving the integration of detailed Aboriginal and Torres Strait Islander data.
- Using business data to link employer characteristics, such as business size, to employee records.

The findings and recommendations in this PIA Update will also assist the ABS and MADIP data custodians in identifying, and assist in managing, the potential privacy impacts of these expected upcoming changes to MADIP.

MADIP is overseen by a Board of senior executive representatives from Australian Government agencies (the MADIP Board). The ABS has undertaken this PIA Update on behalf of the MADIP Board, with independent advice, review and assurance about the PIA process and report provided by external privacy advisors Maddocks. The PIA Update has been conducted to assess compliance with the requirements of the Australian Privacy Principles (APPs), which represent the cornerstone of the privacy protection framework in the *Privacy Act 1988 (Cth)*, and the Office of the Australian Information Commissioner's (OAIC's) *Guide to undertaking privacy impact assessments*. The PIA Update process has also been informed by consultations held from August to September 2019 with a broad range of stakeholder and advocacy groups.

This report provides an assessment of the compliance of MADIP with the APPs, uses the APPs as a framework to analyse additional privacy impacts of MADIP, and makes recommendations to improve APP compliance and privacy best practice. The report also provides information about the scope and conduct of this PIA Update and some contextual information about MADIP, such as project governance and the data and information flows involved in MADIP. The chapters of this report include:

A. Executive Summary (this chapter), which provides a snapshot of MADIP, the PIA Update, and assessment findings, recommendations, and suggestions.
B. Introduction, which looks at MADIP and its project governance in more detail, why this PIA Update was undertaken, the scope of the PIA Update, and how the PIA Update process was conducted.
C. MADIP Data and Information Flows, which takes a deep dive into the data in MADIP and how it moves through the ABS environment.
D. Privacy Impacts and APP Compliance, which assesses the compliance of MADIP (including the new types of data) with the APPs and privacy best practice, and an additional discussion about privacy considerations for the new types of data due to be received into, and linked to other data using the MADIP framework.
E. Next Steps for privacy management in MADIP following this PIA Update.
F. Appendices containing relevant reference material for parts of this report.

1.3 APP Compliance Summary

The following table provides a summary analysis of APP compliance for MADIP and lists recommendations to improve compliance and suggestions to improve privacy best practice. A detailed discussion of each APP can be found in the body of the report.

The recommendations provided in this PIA are intended to apply only to MADIP. In many ways MADIP is a unique and complex project and the advice in this report is designed to assist the ABS and partners to manage MADIP data flows in line with best practice with respect to protecting the privacy of individuals. The best practice suggestions contained in this updated PIA are aimed at enhancing current practices. However, the MADIP Board acknowledges that data custodians (and authorised entities) may implement differing practices to reach the same outcome within their respective agencies.

Table 1 – APP compliance summary, recommendations, and suggestions

APPs	2018 PIA	2019 PIA Update	Commentary	Compliance recommendations	Best practice suggestions
APP 1 **Open and transparent management of personal information**	Partially compliant	Compliant	To improve compliance with APP 1, the ABS published the MADIP Privacy Policy on the ABS website in June 2018. In line with privacy best practice, the ABS has also increased transparency about the data in MADIP and how it is used through online materials which include: • A list of the data in MADIP including potential future data to be linked to MADIP • The legislation under which data are shared for MADIP • A register of ABS data integration projects • A register of MADIP research projects	None	S1. The ABS should update online materials to include more information on: • Non-enduring linkages that have occurred. • Some summary information about the datasets that have been linked to MADIP. • MADIP's governance framework, including a description of the decision and assessment process (including risk assessments) for new linkages and research projects using MADIP. S2. The ABS should continue to obtain advice from a range of sources (including data custodians, whole-of-government building trust initiatives, and stakeholders such as those consulted for this PIA Update) to develop content for online materials for MADIP.
APP 2 **Anonymity and pseudonymity**	Compliant	Compliant	ABS provides some ability in some surveys for anonymity. Otherwise data in MADIP is covered by exceptions to the anonymity principle.	None	None
APP 3 **Collection of solicited personal information**	Partially compliant	Compliant	The ABS is authorised under the *Australian Bureau of Statistics Act 1975* and the *Census and Statistics Act 1905* to undertake surveys and the Census, to collect information (including personal and sensitive information) from other government entities, to link Census data and other information, to produce statistics for analysis, and to publish statistical outputs.	None	S3. The ABS should update online materials to outline the data minimisation approach, including treatment of sensitive data, for MADIP.

APPs	2018 PIA	2019 PIA Update	Commentary	Compliance recommendations	Best practice suggestions
			APP 3 requires that individuals consent to the collection of their sensitive information unless an exception applies. An exception in APP 3 applies to MADIP as the collection by the ABS is authorised by law. In line with privacy best practice, the ABS conducted a review of sensitive data in MADIP which recommended the following improvements to practices: • Minimise data sharing so that only data that are necessary for the purposes of the project are shared and used in MADIP. • Categorised or derived indicators for sensitive data items are used where this is feasible and unless sensitive data items in their original form are required for statistical or analytical purposes. • MADIP project proposals require justification for requesting sensitive data items (including level of detail requested). • Review the retention of sensitive information where there is no compelling business case for retention, or by agreement between ABS and the relevant data custodian. These are in the process of being implemented.		
APP 4 **Dealing with unsolicited personal information**	Compliant	Compliant	Dealing with unsolicited personal information was not an issue for the 2018 PIA, but is now occasionally an issue for MADIP when unsolicited personal information is received by the ABS from data custodians.	None	S4. Before data are disclosed to the ABS for MADIP, the ABS should: • Provide information and assistance to data custodians to aid them in checking data for unsolicited personal information; and

APPs	2018 PIA	2019 PIA Update	Commentary	Compliance recommendations	Best practice suggestions
			The ABS ensures compliance against this APP through the processes it has in place when it collects unsolicited personal information. The Statistical Data Integration Division in ABS has standard operating procedures for dealing with unsolicited personal information for MADIP to ensure the requirements of APP 4 and the *Notifiable Data Breaches* scheme are met. The best practice suggestions for this APP are not required for compliance but rather aim to reduce the risk of unsolicited personal information being shared for MADIP in line with privacy best practice.		• Work with data custodians to decide what reasonable steps, if any, will be taken to reduce the risk of unsolicited information being disclosed to the ABS for MADIP. S5. MADIP data sharing documentation should be updated to highlight the issue of unsolicited information and to note the ABS will work with data custodians to take reasonable steps to reduce the risk of sharing unsolicited personal information with the ABS for MADIP. S6. The ABS should update online materials to provide more information about the management of unsolicited personal information in MADIP.
APP 5 **Notification of the collection of personal information**	Action required	Partially compliant	As a data custodian, the ABS is reviewing and progressively updating privacy collection notices for its household surveys to improve transparency about the potential use of information collected for data integration. A similar update is planned for inclusion in 2021 Census collection notices. Some other MADIP data custodians are increasing transparency about data sharing for MADIP (e.g. The Department of Education has updated its departmental privacy statement to include that personal information is disclosed to the ABS for MADIP.) The ABS published the MADIP Privacy Policy on the ABS website in June 2018.	R1. The ABS should continue to update its APP 5 collection notices, used when the ABS is collecting personal information as a data custodian, to make it clearer to individuals that their personal information may be used for data integration. R2. In relation to collection of personal information by the ABS as the accredited Integrating Authority for MADIP, the ABS should: • Advocate with entities responsible for collection notices to enhance transparency about their disclosure of personal information to the ABS for MADIP by taking reasonable steps to update notices or	S7. Data custodian agency delegates should confirm that, for their disclosure of data to the ABS for MADIP, APP 5 notification requirements have been met and the APP 6 authority for disclosure (and use by the ABS for MADIP) is identified; data custodian delegates may wish to consult with their agency Privacy Officers or otherwise authorised officers for advice on these matters. S8. MADIP data sharing documentation should be updated to provide information that confirms how APP 5 notification requirements are met for collection of data by the ABS for MADIP.

APPs	2018 PIA	2019 PIA Update	Commentary	Compliance recommendations	Best practice suggestions
			Data custodians (or authorised entities) who are APP entities that disclose personal information to the ABS for MADIP need to ensure compliance with APP 5 for their disclosure of information to the ABS. The ABS needs to take some additional steps to cover the personal information it collects from data custodians (or authorised entities) for MADIP by working with data custodians to improve transparency.	otherwise make individuals aware of data use. • Continue to increase transparency about the collection and use of data, including personal information, for MADIP in online materials.	
APP 6 **Use or disclosure of personal information**	Partially compliant	Partially compliant	MADIP data sharing agreements have been updated to require that data custodians identify: • If any data items that will be shared contain personal and sensitive information. • The legislative authority or other for disclosing the data (including any personal information within it) to the ABS. • The reason for sharing data with the ABS. While the existing data sharing documentation for MADIP outlines the legislative or other authority for disclosing data to the ABS, it does not document this in a consistent manner. There is value in a more consistent approach to documenting the authority for disclosure by data custodians to the ABS, particularly where there is no specific legislative authority.	R3. MADIP data sharing documentation should be updated to provide information that confirms the APP 6 authority for sharing data to the ABS for MADIP.	*S7 repeat –* *Data custodian agency delegates should confirm that, for their disclosure of data to the ABS for MADIP, APP 5 notification requirements have been met and the APP 6 authority for disclosure (and use by the ABS for MADIP) is identified; data custodian delegates may wish to consult with their agency Privacy Officers or otherwise authorised officers for advice on these matters.* S9. The ABS should update and maintain online materials that communicate to the public about the data shared in MADIP.
APP 7 **Direct marketing**	Compliant	Compliant	APP 7 is not applicable to MADIP.	None	None

APPs	2018 PIA	2019 PIA Update	Commentary	Compliance recommendations	Best practice suggestions
APP 8 **Cross border disclosure of personal information**	Compliant	Compliant	APP 8 is not applicable to MADIP.	None	None
APP 9 **Adoption, use or disclosure of government related identifiers**	Compliant	Compliant	APP 9 is not applicable to MADIP.	None	None
APP 10 **Quality of personal information**	Compliant	Compliant	Data custodians share personal information to the ABS for MADIP and the quality of this information impacts on the quality of links that can be achieved between datasets. The ABS works with data custodians to improve the quality of information that is shared for MADIP. The ABS has procedures for ensuring the quality of data linkages in MADIP through improving data quality, methodology and infrastructure.	None	None
APP 11 **Security of personal information**	Action required	Action required	In early 2018 the MADIP operating environment was assessed under the Information Security Registered Assessors Program (IRAP). MADIP systems and procedures were certified as compliant against the Information Security Manual by the Australian Signals Directorate. A separate external audit was also conducted in early 2018 on the implementation of functional separation procedures within the Data Linkage Centre (DLC). DLC systems and processes were found to comply with the separation principle. The ABS has a robust framework of legislative, protective security,	R4. The ABS should commit to undertaking a 2 yearly IRAP assessment of the MADIP operating environment as part of a regular program of audits of information security in MADIP. R5. The ABS should finalise and implement the MADIP Data Retention and Destruction policy.	S10. The ABS should enhance information in online materials about the data security protections in place for MADIP.

APPs	2018 PIA	2019 PIA Update	Commentary	Compliance recommendations	Best practice suggestions
			Information and Communication Technology (ICT), and data governance controls for protecting the privacy of individuals and ensuring data security in MADIP. Through this framework, the ABS takes active measures to ensure compliance with APP 11.1 and reasonable steps to comply with APP 11.2. The MADIP Board remains committed to ensuring the security of personal information in MADIP is a core feature of current and potential future data sharing models for MADIP, and will continue to explore the capabilities that future IT environment and infrastructure developments may provide for data security.		
APP 12 **Access to personal information**	Action required	Compliant	The MADIP Privacy Policy now clarifies that: • Individuals can apply to access or correct their personal information held by the agency which originally collected it. • It may not be possible for the ABS to correct or provide access to information in MADIP. Data collected under the *Census and Statistics Act 1905* is subject to legal exemptions to protect the confidentiality of personal information. • As personal information is kept separate from other data in MADIP, it is unlikely that ABS would be able to locate individuals' records to correct them.	None	None

APPs	2018 PIA	2019 PIA Update	Commentary	Compliance recommendations	Best practice suggestions
APP 13 **Correction of personal information**	Compliant	Compliant	As above	None	None
Other suggestions to improve privacy best practice					S11. The MADIP Board should establish and document threshold triggers for future updates of the MADIP PIA. S12. The MADIP Board should consider leveraging government consultation mechanisms to obtain broad-based advice on privacy issues and ethics issues relevant to MADIP.

1.3 Next Steps

MADIP will continue to adapt and evolve to policy and research priorities, methodological and technological advancements, and other environmental changes. New types, and additional years, of data may be linked to MADIP in response to policy priorities and research demands. The ABS is also continuously improving data handling practices and infrastructure for MADIP to preserve privacy, ensure data security, and increase data quality and utility.

This PIA Update is a demonstration of the MADIP Board's commitment to managing the project's privacy impacts. The MADIP Board will continue to take a privacy by design approach for MADIP and acknowledges that another PIA Update or other PIA processes will likely be necessary in the future as MADIP continues to evolve and develop. This report outlines some future developments which are likely to prompt a further update to the MADIP PIA, such as changes to data access or use processes, new types of data proposed for addition, and changes to MADIP's external environment.

The ABS will publish a progress report (similar to the current Implementation Report) on the ABS website within one year of this PIA Update being published to inform on progress of implementing APP compliance recommendations and best practice suggestions.

PART B – INTRODUCTION

The Multi-Agency Data Integration Project (MADIP) is a partnership among Australian Government agencies to develop a secure and enduring approach for combining a broad set of person-centred data from across government domains. Data custodians, or entities authorised by data custodians, share data with the ABS for MADIP. The ABS as the accredited Integrating Authority for MADIP is responsible for receiving, storing, and linking data, as well as assembling extracts of integrated data for analysis and providing access to these extracts to authorised researchers.

The MADIP Privacy Impact Assessment (PIA) is an important part of the governance framework for MADIP and guides MADIP operations. A PIA is a systematic assessment of a project that identifies the impact that it might have on the privacy of individuals, and sets out recommendations for managing, minimising, or eliminating that impact. PIAs are an important component of an organisation's privacy management, and examine not only a project's compliance with privacy legislation, but the broader story of how information is managed within a project.

The last PIA for MADIP was conducted independently by privacy advisory consultants Galexia on behalf of MADIP Agencies and was published on the ABS website alongside the MADIP Agencies' response in April 2018. The ABS is undertaking this update to the MADIP PIA on behalf of the MADIP Board to formally consider developments in MADIP since the last PIA was undertaken and further upcoming changes. The ABS has engaged external privacy advisors Maddocks to provide independent advice and review of the PIA process and PIA Update Report.

This chapter will provide more information about the MADIP project governance framework, the changes to MADIP since the April 2018 PIA and upcoming changes, and the process the ABS has undertaken to conduct this PIA Update and the next steps for privacy in MADIP.

2.1 MADIP

The ABS, Australian Taxation Office, Department of Health, Department of Human Services, and Department of Social Services established MADIP in June 2015 in response to a review of the Commonwealth arrangements for data integration, which recommended enhancing the use of Commonwealth data. From 2015 to 2018, MADIP was in an evaluation phase to test the feasibility and potential value of linking existing public sector data to inform decision making within the government and community for the benefit of the Australian public, consistent with the Australian Government Public Sector Data Management agenda.

MADIP became fully operational under the Data Integration Partnership for Australia (DIPA) in June 2018 and access to MADIP data was provided to authorised researchers in July 2018. MADIP is funded through DIPA until 30 June 2020. DIPA is a three-year $130.8 million investment to maximise the use and value of the Government's data assets, which commenced in July 2017. DIPA is creating new insights into important and complex policy questions through data integration and analysis.

MADIP consists of data assets and infrastructure that are managed within a governance framework. The project governance framework for MADIP and interactions with DIPA are outlined below. More information about MADIP data and information flows and data governance is provided in Part C of this document.

MADIP PROJECT GOVERNANCE

The MADIP Board is the governance body responsible for the operation and strategic direction of MADIP. The Board currently comprises the following agencies as Board members:

- ABS (Chair)
- Australian Taxation Office
- Department of Education
- Department of Health
- Department of Human Services
- Department of Social Services

The Board is also attended by the following agencies as observers and advisers:

- Australian Institute of Health and Welfare
- Department of Industry, Innovation and Science
- Department of the Prime Minister and Cabinet
- Treasury

The governance of MADIP is also guided by the following cross-government committees:

- The DIPA Board – a subgroup of the Deputy Secretaries Data Group, which guides the strategic direction of DIPA and approves funding for data integration research projects that seek to use MADIP in line with policy priorities.
- The MADIP Working Group – a group made up of executive level representatives from the MADIP Board member agencies, which serves the MADIP Board in an issue identification, solution development and advisory capacity.

The Australian Bureau of Statistics (ABS) is the accredited Integrating Authority for MADIP and, in collaboration with its partners, is responsible for combining the data, providing access to authorised users via highly secure ABS systems, and safeguarding privacy—ensuring that no individual person is likely to be identified.

The data custodians involved in MADIP are those authorities responsible for the data included in MADIP. Data custodians are responsible for approving research projects which include the data they responsible for in MADIP. The ABS, as the accredited Integrating Authority, works in partnership with data custodians to ensure that MADIP data are managed in line with ABS legislation, privacy legislation, and relevant custodian legislation. The ABS is also a data custodian in MADIP, such as for the Census of Population and Housing data in MADIP.

Research projects that seek to use MADIP are either initiated through DIPA, or through user funded projects. The DIPA funding model includes an allocation for integrated data research projects. The DIPA Board is responsible for approving the allocation of this funding for projects based on policy priorities. User funded projects must be submitted through a MADIP Board member agency.

For research projects that require new data to be linked to MADIP, the data can be linked to become part of the 'enduring' MADIP asset or as a 'once-off'. Data in the enduring MADIP asset can be made available to authorised researchers for research projects subject to approvals from the ABS and data custodians. Data linked as a 'once-off' is linked for a specific project or projects and accessed by specified authorised researchers and is not retained beyond the completion of the project. However, there may be specific cases to retain data where, during the life of the project, another project is approved involving the same linked data.

The MADIP Board is responsible for approving new enduring linkages to MADIP. In August 2019, the MADIP Board endorsed the MADIP Strategy, which sets the vision and objectives for the project and guides the expansion of the enduring MADIP asset. The MADIP Strategy is provided in Appendix 3.

Relevant data custodians and the ABS are responsible for approving once-off linkages to MADIP and research projects that seek to access existing MADIP data (i.e. projects that don't require new linkage).

The approvals required for linking new data to MADIP and research projects that seek to access MADIP data are summarised in Table 2 below.

Table 2 – MADIP data linkage and access approvals

	The MADIP Board	Relevant data custodians (including the ABS)	The ABS as the accredited Integrating Authority
New enduring linkages	Approval required	Custodians of new data need to approve initial linkage and future updates of their data	Approval required
New once-off linkages		Approval required	Approval required
Data access		Approval required	Approval required

MADIP GOVERNANCE MATERIALS

The governance arrangements for MADIP are reflected in a framework of materials that includes:

- The **MADIP Strategy** – a high-level, future-focused document which establishes a strong, enduring strategic direction for MADIP within the rapidly changing public data landscape.
- The **MADIP Board Terms of Reference** – outlines the role, responsibilities, and operation of the MADIP Board.
- The **MADIP PIA** – identifies potential impacts MADIP may have on the privacy of individuals, and sets out recommendations for managing, minimising, or eliminating these impacts.
- The **MADIP Operating Model** – reflects current practices and describes technical and procedural details relating to data sharing and use in MADIP.
- **Data sharing documentation** – forms a record of the specifications and terms of sharing data with the ABS for MADIP and facilitates data sharing within relevant legislation.
- ABS **Data Integration Plans** – document how ABS meets its commitments as an accredited Integrating Authority including project risk assessments, its legislative and privacy obligations, and internal policies for each linkage project.
- **Research project proposals** – capture project details to facilitate approvals from relevant data custodians, including the ABS, and ensure access to data are conducted within the Five Safes Framework.

2.2 PIA Update – Purpose, Scope and Approach

The primary purpose of this PIA Update is to consider the developments in MADIP since the 2018 PIA was undertaken, as well as new types of data planned to be linked to MADIP, and:

- Analyse the impacts that changes to the MADIP since the 2018 PIA will have on the privacy of individuals whose personal information is collected by the ABS and whose unidentified information is made available to authorised researchers;
- Identify privacy risk areas in relation to compliance with the APPs and community expectations; and
- Identify, assess, and where appropriate recommend, options for managing negative privacy impacts to improve compliance against the APPs and make suggestions in line with privacy best practice.

MADIP DEVELOPMENTS SINCE APRIL 2018 AND FUTURE NEW DATA

MADIP is an evolving project. New types and additional years of data are linked through MADIP in response to policy priorities and research demand. The ABS is also continuously improving data handling practices and infrastructure for MADIP to preserve privacy, ensure data security, and increase data quality, timeliness and utility.

There have been a number of developments in MADIP since the 2018 PIA was undertaken:

- The amount of datasets linked through the MADIP environment has expanded, with additional years and new sources of data of the same broad type having been collected, integrated and stored using the MADIP framework.
- ABS survey data has been collected, integrated and stored (which was undertaken after a specific PIA was conducted using the MADIP framework).
- The ABS is using a new data system/environment for securely receiving, storing and linking MADIP data.
- The ABS has enhanced its person-centred data linkage infrastructure to more efficiently combine datasets on a project by project basis and strengthen privacy protections.
- The process for providing authorised researchers with access to MADIP data for research projects has been reformed and streamlined to deliver more timely access in a secure way.
- More detailed information about MADIP is available on the ABS website.

These developments have not yet necessitated an updated or new MADIP PIA based on PIA threshold guidelines from the OAIC. This is because the data linked through MADIP since the 2018 PIA have been of the same broad types of data that the 2018 PIA assessed. All new data linked to MADIP have been subject to internal ABS privacy risk assessments. An exception here has been the linkage of the 2014-15 National Health Survey to MADIP, which was supported by a separate independent PIA. Also, the developments in MADIP since the 2018 PIA have not yet resulted in a significant change to the handling of personal information in MADIP.

However, there are some upcoming changes to MADIP that will involve integrating different types of data:

- Non-sensitive ABS household survey data[1]
- Detailed Aboriginal and Torres Strait Islander data
- Selected business characteristics

This PIA Update has been conducted to support the inclusion of these new types of data in MADIP. The PIA Update is also an opportunity to formally examine the developments in MADIP since the 2018 PIA was undertaken.

The reforms to data access are described below. The other developments in MADIP since the 2018 PIA and the upcoming new types of data listed are described in the MADIP DATA AND INFORMATION FLOWS chapter of this document.

ABS Data Access Reforms

In 2019, access to ABS microdata has been streamlined through the ABS Data Access Reforms initiative taking into account feedback from researchers and data custodians. This initiative has enabled more researchers to access MADIP microdata faster and more easily in a safe and effective way.

Previously, the MADIP Basic Longitudinal Extracts were the only extracts available for access by authorised government and non-government researchers via non-secondment arrangements. Now, government employees, government contractors, and individuals sponsored by government departments can apply to access a range of custom MADIP microdata extracts under non-secondment arrangements. The ABS is also reviewing the 'Safe People' element of the Five Safes Framework to facilitate access to ABS microdata for other non-government researchers (particularly academics and researchers from public policy research institutes).

Access continues to be managed within the Five Safes Framework, and the ABS's legislative framework (i.e. the data must be unidentified and the information is disclosed in a manner that is not likely to enable the identification of an individual). As with current MADIP research projects, access will be subject to ABS and data custodian approval.

THE SCOPE OF THE PIA UPDATE

The MADIP PIA Update identifies and assists in managing the potential privacy impacts of developments in MADIP since the 2018 PIA in the context of current and expected sharing and use of data for MADIP.

The PIA Update describes the data and information flows in MADIP and the current infrastructure and governance arrangements that determine how MADIP data are managed. The new types of data that are due to be included in MADIP will be highlighted and described in more detail.

The PIA Update also assesses the compliance of MADIP against the APPs. Based on this assessment, the PIA Update makes recommendations for improving compliance and also steps that can be taken for improving privacy best practice.

MADIP has standard processes for data flows and projects, so the assessment will mostly focus on MADIP as a data system. Where relevant, particular impacts that certain data types have on APP compliance and forming recommendations will be considered as part of the broader assessment. The recommendations and suggestions made through this PIA Update will be targeted at an issue level to provide a framework for the future inclusion of similar data in MADIP.

The scope of the PIA Update is limited to the data integration activities of MADIP. For the purposes of this document, 'MADIP data integration activities':

[1] Survey data that would not be considered sensitive information when the data includes personal information.

- Are conducted within the MADIP governance framework;
- Use the person linkage spine; and
- Involve at least one enduring MADIP analytical dataset.

The PIA Update focuses on the sharing and use of data (including personal information) throughout the linkage process, from when the ABS receives the data through to statistical outputs being cleared by the ABS team for dissemination. The recommendations made and best practice steps suggested in this PIA Update require action by the ABS as the accredited Integrating Authority as well as by data custodians (including the ABS) that are involved in MADIP. The recommendations and suggestions made herein are specific to MADIP, the MADIP Board, MADIP data custodians and the ABS as the accredited Integrating Authority for MADIP; they do not extend beyond MADIP to these parties' other operations, or any other parties.

The PIA Update does not:

- Cover the data activities of ABS or MADIP data custodians as a whole;
- Assess the application of, or compliance with, restrictions on the handling of information which is protected pursuant to secrecy provisions in portfolio legislation (other than the extent to which the collection, use, or disclosure of personal information will be 'authorised or required by law' pursuant to those secrecy provisions for the purposes of APPs 3.4, 3.6 and 6.2);
- Cover personal information that is not collected for data integration but may be collected and handled in the context of the project (e.g. personal information about researchers who apply for data access); or
- Consider compliance by researchers that access MADIP data with applicable privacy laws or secrecy provisions.

THE APPROACH TO UPDATING THE PIA

The approach taken to updating the MADIP PIA follows the *OAIC Guide to undertaking privacy impact assessments* and uses its 10 steps to undertaking a PIA. The PIA Update was conducted by the ABS on behalf of the MADIP Board, with privacy advice and independent review and assurance provided by external privacy advisors Maddocks. Maddocks was selected to provide this review and assurance following a competitive selection and procurement process.

The 2018 PIA was undertaken by Galexia, independently from the ABS and the MADIP Board. This independent PIA assisted MADIP in moving from an evaluation phase into production, and further built on the standard of robust privacy management in the project. The MADIP Board agreed to the ABS conducting the MADIP PIA Update and the involvement of independent advice and review as being an efficient and effective way to update the MADIP PIA and build the privacy capability of MADIP agencies whilst ensuring a robust and transparent process.

PIA Update Report

In preparing the PIA Update Report, the ABS:

- Prepared a project plan for the conduct of the PIA Update, including defining scope;
- Procured the services of the independent advice, review and assurance provider (Maddocks);
- Identified and examined the changes to MADIP since the 2018 PIA;
- Sought expert advice from Maddocks on the ABS' planned review of current MADIP practices, consultation and strategy and the PIA Update process;
- Undertook a broad consultation to gauge targeted stakeholder views on MADIP and data integration and prepared a summary of the consultations;
- Prepared initial recommendations for compliance and best practice for consideration by the MADIP Board;
- Prepared the PIA Update Report with advice from consultation sessions, Maddocks, the MADIP Board; and
- Published the PIA Update Report.

MADIP Board Response

The MADIP Board will develop a response to this Report which will also be published on the ABS website.

Independent Review Report

Maddocks will provide independent review and assurance in relation to this Report. The Maddocks report will be published on the ABS website when it becomes available.

Implementation Report

The MADIP Board will publish a progress report (similar to the current Implementation Report) on the ABS website within one year of this PIA Update being published to inform on progress of implementing recommendations and best practice suggestions from the PIA Update process.

CONSULTATION

Consultation was an essential part of conducting this PIA Update and an opportunity to hear stakeholder views on MADIP and its privacy management arrangements. A summary of consultation outcomes is available on the ABS website.

Stakeholders were selected from academic, civil society, government, privacy and Aboriginal and Torres Strait Islander sectors, aiming to be representative of the broad range of stakeholders with an interest in MADIP. Some groups consulted were part of the 2018 PIA consultation while some were added to augment the representation of the privacy sector and particular population groups. A list of parties consulted is included in Appendix 4.

The MADIP Board was engaged in the PIA Update via Board meetings and out-of-session communication. The consultation approach was endorsed by the MADIP Board and Maddocks.

To consider the proposal to link detailed Aboriginal and Torres Strait Islander data to MADIP, ABS conducted two targeted consultation sessions with Aboriginal and Torres Strait Islander stakeholders, in conjunction with the ABS' Centre of Excellence for Aboriginal and Torres Strait Islander Statistics. The ABS engaged with an existing group that advises the 2021 Census (the Census 2021 Remote Expert Review Panel) and the ABS Round Table on Aboriginal and Torres Strait Islander Statistics (which has Australia-wide representation). The aim of these sessions was to provide information about MADIP and to hear any views and concerns, particularly around the proposed linkage of detailed Aboriginal and Torres Strait Islander data to MADIP.

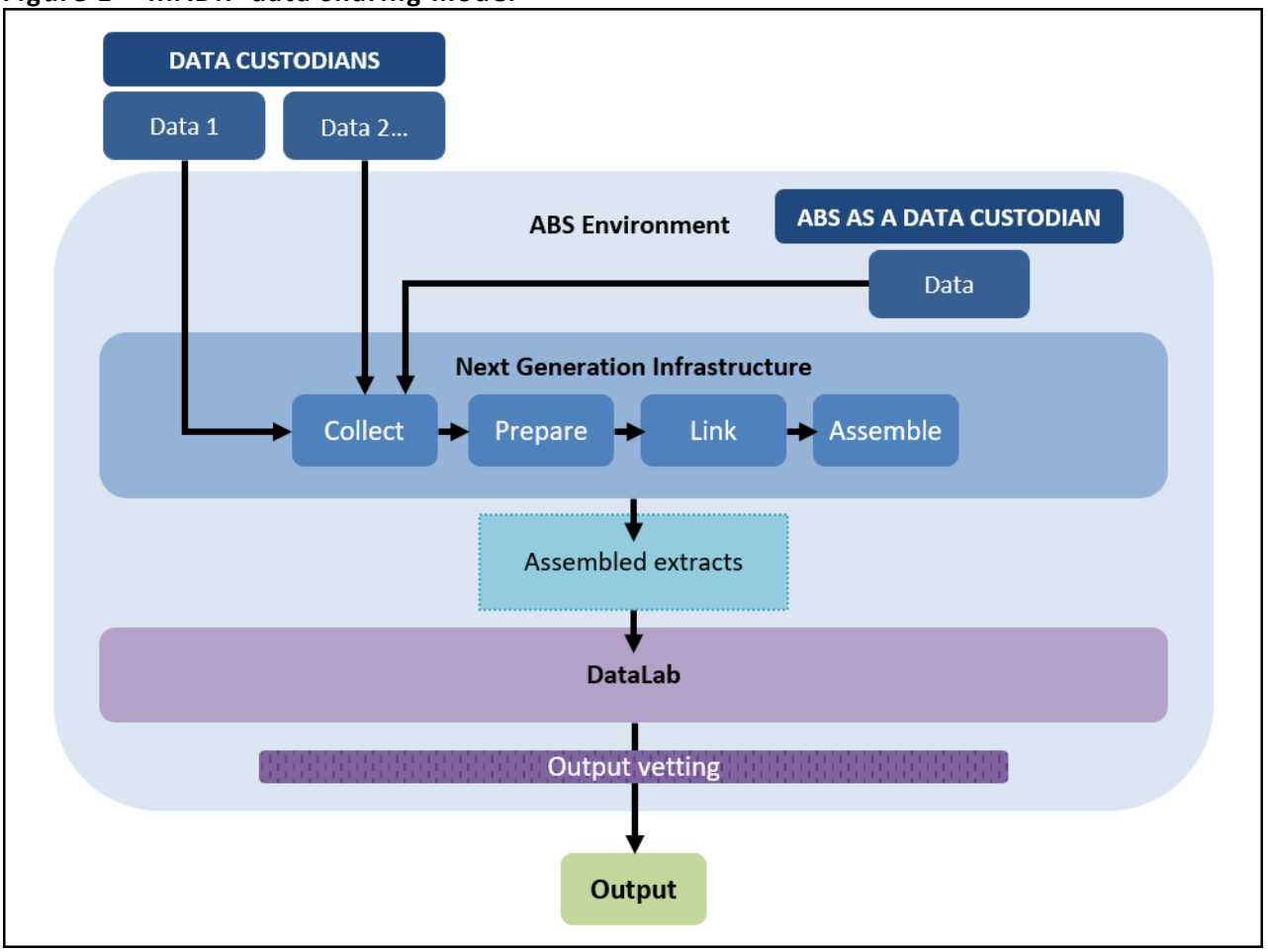
Relevant information from recent privacy and trust consultations was considered. The ABS took into account other consultation that is occurring, such as ABS consultation for the 2021 Census independent PIA and the consultation by the Office of the National Data Commissioner for the proposed Data Sharing and Release legislation PIA.

PART C – MADIP DATA AND INFORMATION FLOWS

MADIP is a system of governance and data infrastructure that enables the secure integration of a broad set of person-centred data for statistical and research purposes. MADIP has established an enduring, longitudinal, multi-topic integrated data asset. MADIP also allows for the integration of datasets for specific projects that do not become part of the enduring asset.

The data sharing model for MADIP is centred on the ABS as the accredited Integrating Authority for the project. Data custodians and entities authorised by data custodians share data with the ABS for MADIP. The ABS links the individual datasets supplied for MADIP to a central linkage infrastructure, assembles analytical extracts of linked datasets for specific purposes, and provides secure access to these extracts for authorised researchers for approved projects. The ABS does not share MADIP microdata outside of the ABS environment. This data sharing model is pictured in Figure 1 below.

Figure 1 – MADIP data sharing model



This chapter describes the data in MADIP and how information flows through the ABS. MADIP is conducted with a privacy by design approach which incorporates consideration and management of privacy risks into each step of the data integration process, from data collection through to dissemination and access.

This chapter also outlines:

- The personal and sensitive information in MADIP and how that information is treated; and
- The governance arrangements that support data sharing and use in MADIP.

3.1 Personal Information

MADIP includes personal information such as name, address, and date of birth. This information is used to build the central linkage infrastructure and to link separate datasets to this infrastructure. Direct identifiers are stored separately from other information in MADIP in accordance with the separation principle. This other information may, in some circumstances, be considered personal information even when it is separated from direct identifiers as it may enable the re-identification of an

individual (e.g. through the combination of data items). Access to personal information in MADIP is strictly controlled and limited to a small team of ABS staff.

The MADIP data that the ABS makes available for authorised researchers in the secure ABS DataLab does not include personal information as it 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 1988 (Cth)*[2]). The ABS uses the Five Safes Framework to manage disclosure risks associated with providing access to de-identified MADIP data.

MADIP also includes 'sensitive data' (i.e. data that contains information that fits into the categories of sensitive information defined in the *Privacy Act 1988 (Cth)* even when the data does not include personal information). This includes:

- Health information;
- Ethnicity and racial background;
- Indigenous status[3];
- Religious affiliation; and
- Sexuality[4].

The ABS is implementing some additional principles for managing this sensitive data in MADIP. These principles will apply even when data do not contain personal information, and are as follows:

- Data sharing is minimised so that only data that are necessary for the purposes of the project are shared and used in MADIP.
- Categorised or derived indicators for sensitive data items are used where this is feasible unless sensitive data items in their original form are required for statistical or analytical purposes.
- MADIP project proposals require justification for requesting sensitive data items (including level of detail requested).
- Sensitive data are destroyed where there is no compelling business case for retention, or by agreement between the ABS and relevant data custodians.

Regardless of whether the data contains personal information or not, the ABS treats all MADIP data in the ABS environment with standards appropriate for personal information.

The Information Flows section below provides more detail on the flow of information through the ABS environment for MADIP.

3.2 Data Governance

The management of MADIP data is determined by arrangements agreed between data custodians and the ABS as the accredited Integrating Authority. These arrangements are documented in the following governance materials, which are described in more detail below this list:

- The MADIP Operating Model
- Data sharing documentation
- ABS Data Integration Plans
- Research project proposals

The MADIP Strategy and the MADIP PIA provide a framework for these governance materials.

[2] 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).
[3] Indigenous status is not a category of sensitive information in the *Privacy Act 1988 (Cth)* but is considered a type of information about ethnicity and racial background.
[4] Information on sexuality is not currently contained in datasets collected and stored in MADIP, but it could be inferred through other information.

MADIP Operating Model

The MADIP Operating Model reflects current practices and captures technical and procedural details relating to data sharing and use in MADIP, as agreed by the MADIP Board. It also:

- Details standard processes used throughout the end-to-end journey for MADIP data including supplying data to the ABS, data linking processes and model, updating and maintaining the enduring MADIP asset, creating assembled linked data extracts, data access, confidentiality, communication, data retention and destruction, and breach management; and
- Defines the responsibilities of data custodians, including the ABS, and the ABS as the accredited Integrating Authority, and also the roles of the MADIP Board and DIPA Board.

Data sharing documentation

Data sharing for MADIP is supported by governance documentation that usually includes one or more of the following depending on the requirements of the ABS and data custodians:

- Memorandum of Understanding
- Letter of Agreement
- Public Interest Certificate
- Other official letter or email from an authorised delegate

The primary purpose of this data sharing governance documentation is to ensure data sharing occurs in line with the legislative requirements of the relevant data custodian and the ABS. This documentation may also provide a record of the terms of data use agreed between data custodians and the ABS.

ABS Data Integration Plans

Data Integration Plans support safe data integration projects that meet the ABS' accredited Integrating Authority responsibilities and ensure all projects are of benefit to the public. These plans use a standard template that includes information on:

- The project purpose and public benefit;
- Approvals from relevant data custodians and project owners including ABS Senior Executive approval for every project;
- Details of datasets to be used;
- Summaries of the linkage strategy, access strategy, data retention strategy, proposed outputs, and adherence to ABS data management processes (e.g. the separation principle);
- The data information flows for the project;
- Legislative or other authority for sharing the data to the ABS;
- Compliance with relevant legislation for using the data;
- A PIA threshold assessment and links to a PIA where this is required;
- Risk assessments based on the Commonwealth Arrangements, or for projects out of scope of these arrangements – on the Five Safes Framework.

Research project proposals

MADIP research project proposals are standard forms developed by the ABS to capture project details, facilitate approvals from relevant data custodians, and ensure access to data is conducted within the Five Safes Framework. The proposal form requests details from project proponents on the following:

- Research objective and methodology for the project;
- Data required, with further details about how data will be used for data with particular legislative requirements (e.g. Social Security and Related Information data);

- Whether the project has been submitted for ethics approval or grant review;
- Names and positions of the researchers who will get access to data and also any other project team members who the analysis will be discussed with but who will not get access to unvetted data; and
- Anticipated outputs and dissemination approach.

3.3 The Data in MADIP

MADIP brings together a broad set of person-centred data from across government domains to provide whole-of-life insights that can improve the lives of all Australians and to facilitate the use and re-use of public data for statistical and research purposes.

At the time of the 2018 PIA, MADIP contained Census data and a number of Commonwealth administrative datasets. Since then, the volume of data in MADIP has increased significantly. The increase in data includes some ABS household survey data that has been supported by a separate PIA process.

Commonwealth data still makes up most of the data in MADIP, but some linkages with state government data have occurred since the 2018 PIA.

MADIP contains data on a broad range of topics, including healthcare, education, government payments, personal income tax, and population demographics (including the Census). The data in MADIP also stretches over a number of years; for example, migration data from 2000 to 2019 is included in MADIP.

The broad range of data sources in MADIP allows information from different datasets to be analysed together to unlock new insights for different groups in Australia. For example, the data can be used to analyse how the background characteristics and living conditions of vulnerable groups relate to their use of welfare and medical services to design and target these services more effectively.

The longitudinal nature of MADIP data allows changes and patterns in the Australian population to be better understood and analysed across time. For example, the data can be used to analyse how family background and different educational choices influence post-school education and employment outcomes for students, to inform government support payments and programs for students that aim to improve educational access and outcomes.

There are three ways that data are added to MADIP:

- To form the central linkage infrastructure – called the 'person linkage spine';
- As part of the enduring analytical asset; or
- As a once-off linkage.

The person linkage spine used for MADIP is currently made up of Medicare Enrolments Database, Personal Income Tax, and Social Security and Related Information data. This infrastructure is described in more detail in the Information Flows section below.

The enduring MADIP asset contains data sources that can answer a broad set of policy and research questions and is maintained and made available for multiple approved research projects. The MADIP Strategy defines a vision, a set of objectives, and a list of benefits for the project that the MADIP Board considers when deciding to enhance and maintain the enduring MADIP asset.

Data linked as a once-off is linked for a specific research project, or projects, and set of authorised researchers and is not retained beyond the completion of the research project. However, there may be specific cases to retain data where, during the life of the project, another project is approved involving the same linked data.

Appendix 5 provides a full list of the datasets in MADIP as well as the following information on the data:

- The name of the dataset and the reference period included in MADIP;
- The data custodian;
- A short description of the data;

- The legislative or other authority for sharing the data to the ABS for MADIP;
- Whether the data includes data items that would be considered sensitive information if the data includes personal information;
- Whether the data has been linked into the enduring analytical asset or as a once-off; and
- Whether the data was included in MADIP for a DIPA project or not.

The table in Appendix 5 also lists upcoming datasets that are expected for inclusion in MADIP and a limited set of the above information where this is available. A list of the current and upcoming data in MADIP is also available on the ABS website.

3.4 New types of data in MADIP

This PIA Update has been conducted to support the inclusion of new types of data in MADIP. These data types are non-sensitive ABS survey data, detailed Aboriginal and Torres Strait Islander data, and a small number of business characteristics. The stakeholder consultation undertaken to inform this PIA Update included specific discussion of these new data types in MADIP.

The inclusion of these new types of data in MADIP is being driven by DIPA research projects, which are also described below. This PIA Update sets a framework for the future inclusion of non-sensitive ABS survey data and a small number of business characteristics beyond these research projects.

A discussion of some additional privacy considerations for these new data types is provided in the New Data Types – Discussion section of this document.

Survey data (non-sensitive)

The ABS runs a large number of surveys on a broad range of topics to collect information from households and businesses to produce official statistics on a wide range of economic, social, population, and environmental matters of importance to Australia. Data from ABS household surveys in particular are a rich source of information that can lead to new insights when integrated with other data sources, such as through MADIP.

In contrast to the administrative data that makes up much of the MADIP asset, the ABS is the data custodian of survey data, which it directly collects from individuals. Where approved and following standard governance arrangements, survey data is shared internally within the ABS for MADIP. Once shared internally, ABS survey data follows the standard information flows for MADIP described in the next section.

This PIA Update considers the inclusion of non-sensitive survey data in MADIP. I.e. survey data that would not be considered sensitive information when the data includes personal information. The integration of sensitive survey data to MADIP needs to be supported by a separate PIA process.

The enduring MADIP asset already includes sensitive data from the National Health Survey (NHS) 2014-15, which was covered by its own independent PIA. Data linkage activities have been completed for the inclusion of the Survey of Disability, Ageing and Carers (SDAC) 2018, which also includes sensitive data, in MADIP. The independent PIA for the NHS 2014-15 linkage to MADIP also provides a framework for the SDAC linkage. The PIA Update also provides a useful framework for including SDAC 2018 in MADIP, so access to this data will commence after this PIA Update is complete to support two priority approved DIPA projects. These projects are summarised below:

The First Five Years: What makes a difference?

Led by the Department of Education, this project will enhance understanding of the effects of health and socio-economic factors that drive disadvantage with respect to children's early development outcomes. The inclusion of SDAC 2018 data will help identify particular early childhood policy interventions or protective factors that can improve these outcomes.

Health Mind, Healthy Body

Led by the Department of Health, this project aims to identify policy opportunities to improve life expectancy and health outcomes for people with mental health diagnoses and co-morbid physical health conditions. The SDAC 2018 data will support

analysis to ensure that people affected by mental health conditions are being appropriately supported by the social welfare system.

The demand for MADIP data is growing and many MADIP data users recognise the value of ABS survey data, so the ABS expects that more projects will seek to link ABS survey data to MADIP in the future.

Detailed Aboriginal and Torres Strait Islander data

MADIP contains Aboriginal and Torres Strait Islander data in the form of Indigenous status data items from datasets in MADIP including the Census, Social Security and Related Information, and Australian Apprenticeships Incentives Program & Training Contracts. For example, the 2016 Census question that collected information on Indigenous status asked respondents if they were of Aboriginal and/or Torres Strait Islander origin.

There is a 2019/20 DIPA project that proposes to include more detailed Aboriginal and Torres Strait Islander data in MADIP in order to understand how people are faring after they cease participating in employment services programs in remote Australia. Led by the Department of the Prime Minister and Cabinet and the National Indigenous Australians Agency (NIAA), this project aims to discover whether participants have gained sustainable employment, experienced different outcomes such as moving to regional and metropolitan areas to look for work, or if they are no longer in the labour force.

In order to achieve this, the project proposes to use data from the NIAA relating to the following programs:

- Remote Jobs and Communities Program
- Community Development Program
- Indigenous Employment Program

It is intended that the project will link NIAA program data with the following MADIP datasets:

- Census of Population and Housing
- Australian Apprenticeships Incentive Program and Training Contracts
- Social Security and Related Information
- Personal Income Tax
- Medicare Enrolments Database
- Medicare Benefits Schedule
- Pharmaceutical Benefits Schedule

This PIA Update has not assessed the use of detailed Aboriginal and Torres Strait Islander data in MADIP beyond this DIPA project. Future research projects that seek to use detailed Aboriginal and Torres Strait Islander data in MADIP may need to be supported by a separate PIA process.

High level business characteristics

MADIP currently only includes person-centred data that comes from information collected from persons and households. However, the ABS and many government and research stakeholders recognise the significant benefits that would come from combining person-centred data with business data.

The ABS maintains another large integrated data asset called the Business Longitudinal Analysis Data Environment (BLADE). BLADE combines business tax data, government program data, and information from ABS surveys over time to provide a better understanding of Australian businesses and the economy.

There are two 2019/20 DIPA projects that are proposing to link selected business characteristics such as employer size, industry and turnover to employee records MADIP. These projects are summarised below:

Migration's Impact on Australian Society

Led by the Department of the Prime Minister and Cabinet, Treasury, Department of Home Affairs, and the ABS, this project aims to evaluate various impacts of international migration on Australia. The inclusion of broad business information with migration data currently in MADIP would support the analysis of the characteristics of businesses that benefit from employer-sponsored migration.

Exits from Income Support

Led by the Department of Social Services, the Department of Education, the Department of Employment, Skills, Small and Family Business, the Department of Health, the ABS, and the Treasury, this project aims to build a broader picture of income-support recipients' pathways once they leave the system, and if possible, the interventions that led up to this outcome. The addition of broad business information would support the analysis of the types of businesses employing former recipients of income support.

3.5 Information Flows

This section describes how the information and data in MADIP flows through the ABS environment. It provides information on:

- The data environment and a core principle of data integration – the separation principle;
- Data collection and preparation;
- Data linkage;
- Interoperability with other integrated data assets;
- Assembling linked data extracts;
- Data access and the 'Protari' Application Programming Interface;
- Releasing outputs; and
- The additional procedures the ABS uses for managing Census information.

The MADIP information flows are also depicted in a diagram in Appendix 6.

The current information flows for MADIP are not significantly different to those described in the 2018 PIA. The key changes are small improvements to methods and processes, and interoperability with other integrated data assets.

The separation principle and secure data environment

All data integration for MADIP is done in accordance with the separation principle[5], which means that no individual can access both the identifying information used for linkage, such as name, address and date of birth, together with the analytical information which does not contain direct identifiers. The ABS implements the separation principle through a 'functional separation' approach. Functional separation is a collection of access controls and procedures to restrict and regulate access to data. It involves the use of discrete 'functional roles':

- Librarian: Prepares, standardises, and anonymises identifying data used for linkage. Typically the linkage data are comprised of variables relating to name, address, date of birth, and sex or gender.
- Linker: Links datasets using anonymised linkage data
- Assembler: Uses the linkage results to create linked analytical data.

The two key tenets of functional separation are that an individual may never occupy more than one functional role at any given time, and that data passes progressively through the functions.

All Librarian, Linker, and Assembler activity occurs in an isolated IT environment – the Data Integration Next Gen Infrastructure (NGI). This environment has no external connectivity (i.e. no email, internet, etc.) to mitigate risk of data exfiltration. Baseline

[5] https://statistical-data-integration.govspace.gov.au/topics/applying-the-separation-principle

security clearance is required to obtain access to the DI NGI. Each functional role has separate access controlled data holdings within the DI NGI.

Data collection and preparation

When data are supplied to the Librarian team for MADIP, whether that be from ABS or external data custodians, best practice is that the linkage and analytical variables are received as separate files. On rare occasions when a data custodian (or entity authorised by a data custodian) is unable to supply data in this way, an officer in the Librarian functional role will split the data as soon as it is received.

Files are acquired into the ABS corporate environment, either through online secure file transfer portals, or encrypted physical media (e.g. USB devices), and securely transferred into the DI NGI. Identifying data are transferred to the Librarian holdings and analytical data are transferred to the Assembler holdings.

Librarian staff clean, standardise, and anonymise linkage data before securely transferring it to Linker staff. Librarians anonymise names using character substitution, and standardise date of birth and sex or gender into standard formats. Librarians also geocode addresses to an Address Register ID (ARID), as well as Mesh Blocks and higher geocodes from the Australian Statistical Geography Standard (ASGS). An ARID is an identifier representing a unique Australian address. Librarians anonymise ARIDs before transferring them to Linkers so that their corresponding address cannot be 'looked up' on the Address Register.

Data linkage

Linker staff link datasets using the anonymised linkage data. Medicare Enrolments Database (MEDB), Social Security and Related Information (SSRI), and Personal Income Tax (PIT) data are linked to create the Person Linkage Spine ('the spine'). The spine is the result of a three-way linkage between these datasets (the 'core' spine datasets) and is the linking infrastructure that supports all MADIP linkages. Persons who are present on at least one of these three data sources are included on the spine. The spine itself is a concordance or 'map' of the links between the core spine datasets. To link new data to the MADIP data asset, the required linkage variables from the core spine datasets are brought together with the spine, and the linkage variables from the new dataset. This new data can be ABS data or data provided by external data custodians or entities authorised by data custodians. The result of a new linkage is a 'linkage results file', which, like the spine, is simply a concordance or map between the IDs on the spine and the new linked dataset. No linkage variables are included on the linkage results file.

Interoperability

While the ABS is the accredited Integrating Authority responsible for the MADIP asset, the Australian Institute of Health and Welfare (AIHW) is also an accredited Integrating Authority. As part of the DIPA program, the ABS and AIHW are undertaking a project to identify methods and establish governance arrangements to enable these agencies to combine integrated data created and held separately by each agency.

One demonstration project has been completed (the Older Australians Project) and another is currently underway (the Victorian Community Health Project).

To progress these projects, the AIHW linked administrative data to its person linkage spine (based on the MEDB dataset), and shared the results with the ABS. The ABS acquired the analytical data from the data custodian and combined it with analytical variables from MADIP. Essentially this meant that the librarian and linker functions were undertaken at the AIHW, while the assembly activity was undertaken at the ABS. No personal information collected and stored by the ABS or the AIHW was disclosed between the two agencies.

The benefits of this approach include increased efficiency and timeliness for analysts, a reduction in duplication of effort between ABS and AIHW (in that the two agencies are not undertaking identical linkage work), and data minimisation. The last point is particularly relevant for this PIA, in that sharing the results of linkage work means that the ABS does not need to acquire personally identifying information from data custodians that has already been provided to the AIHW. Instead of acquiring name, address, sex, and date of birth information from the data custodian for linkage purposes, the ABS only needs to acquire the linkage results from AIHW (in the form of securely transferred and de-identified linkage keys).

In the Older Australians Project, the linkage keys were used to identify a common population that enabled analysis of one dataset containing information on use of aged-care services from AIHW to be directly compared with a different MADIP dataset from the ABS. In the Victorian Community Health Project, the linkage keys will be used to create a common dataset that combines MADIP with Victorian Department of Health and Human Services (DHHS) data. The diagram provided in Appendix 7 shows the data flows in the Older Australians Project, which could be adapted to a variety of settings with other data custodians.

Data assembly

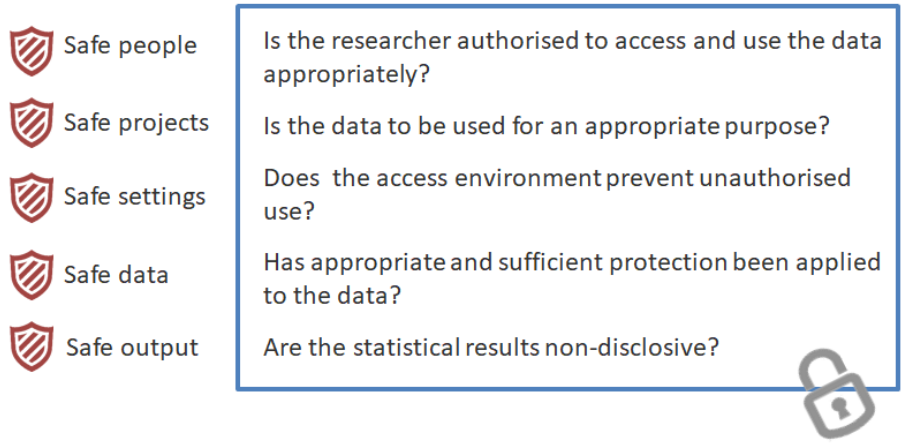
Following linkage, the spine IDs and the linkage results are securely transferred to the Assembler holdings. Personal identifiers are not transferred as part of this process. The Assembly team uses the transferred information to bring analytical variables from different datasets together into an assembled analytical file. This is conducted in line with conditions of use set by data custodians (in the vast majority of cases this process refers to linkage conducted by and within the ABS but it is also relevant to the data flow from the interoperability demonstration projects where linkage is conducted by the AIHW). Additional checks are applied to ensure that the IDs on the assembled analytical file are replaced with analytical IDs before it is transferred out of the DI NGI and made available to approved analysts through the ABS DataLab. Analytical variables are also checked to ensure that any data that could create a credible risk of data users spontaneously recognising individuals are detected and treated to mitigate the risk before it is released for analysis.

Data access

The ABS provides access to MADIP data to authorised researchers in the highly secure ABS DataLab. The DataLab is a data analysis solution for high-end users who want to undertake interactive (real-time) complex analysis of microdata. The ABS manages access to MADIP data by using the Five Safes Framework – an internationally recognised approach to managing disclosure risk summarised in Figure 2 below and described in more detail in APP 11. Access to MADIP data in the DataLab for a research project is subject to ABS and data custodian approval.

Consistent with the Australian Government's Secure Cloud Strategy, the ABS is considering plans to use secure cloud storage and computing for the DataLab (the 'Cloud DataLab Project'). All data will continue to be securely stored within Australia and managed so information is not released in a manner that is likely to enable the identification of an individual. The Cloud DataLab Project is not in scope of this PIA Update and will be supported by a separate PIA process to be conducted by the ABS.

Figure 2 – The Five Safes Framework



There are two types of MADIP microdata products that researchers can access:

- MADIP Basic Longitudinal Extracts – microdata assembled from a pre-prepared file by the ABS that contain key demographic, social, healthcare, education, government payment and income information; and
- Custom MADIP extracts – microdata extracts built or reused according to researcher specifications.

Access to MADIP microdata is enabled by non-secondment arrangements, via Section 15 of the *Census and Statistics (Information Release and Access) Determination 2018*. Under this legislation, all data must be unidentified and the information is disclosed in a manner that is not likely to enable the identification of an individual.

There may be a small number of cases where government secondment is required, such as self-assembly of custom extracts.

The MADIP Basic Longitudinal Extracts are available for submission of project proposals by authorised government and non-government researchers. An outline of these extracts (including a data item list and test file) is available from the ABS website.

Custom MADIP microdata extracts are available for submission of project proposals by authorised government employees, government contractors and individuals sponsored by government. A list of existing custom MADIP microdata extracts for reuse is available on request from the ABS.

The ABS is reviewing the Safe People Framework, and is considering whether additional and strengthened controls could be put in place to enable access to academics and researchers from public policy research institutes not associated with government.

Data access via 'Protari'

After a 'standard' release into the DataLab of MADIP microdata (e.g. one advertised on the ABS website), the same data may be released via the Protari Application Programming Interface (API). Whereas the DataLab enables access to the actual microdata, Protari enables users to create aggregate tables that are confidentialised on the fly without ever accessing or seeing the underlying data. In the future, customised microdata from MADIP could also be released via Protari. The confidentiality of Protari outputs is ensured by the same perturbation methods that are used in ABS TableBuilder, which is used extensively for secure Census and survey dissemination. The ABS is running a trial of the Protari API to assess the utility of its API functionality.

Releasing outputs

The ABS checks all data outputs produced by researchers in the DataLab that are intended for release before these outputs leave the DataLab. Only aggregated information can leave the DataLab. Requests to release unit records are refused.

All outputs are vetted and if necessary treated by specialised ABS staff. This process controls the risk of individuals being re-identified through aggregate outputs, such as through unusual combinations of data items. The output vetting process assumes that the outputs will be publicly released, even if that is not the researcher's intention.

Management of Census information

The data flow for integration of 2016 Census data included extra functional roles in order to meet ABS commitments about use of Census data, particularly names. Librarian work for Census 2016 name data was done in a separate 'Census Name Manager' functional role, and the Census names were irreversibly encoded using a method called 'lossy encoding'[6] before being transferred to Linker holdings.

2011 Census data were integrated to MADIP with a lower quality linkage that did not use name or full address, as they had not been retained, so the provisions used for 2016 Census data were not necessary for 2011 Census data.

The data flows for the integration of 2021 Census data to MADIP are still under consideration and will be informed by the 2021 Census PIA and reflected in a future update to the MADIP PIA (or a separate PIA).

[6] https://www.abs.gov.au/websitedbs/d3310114.nsf/home/Information+paper+Name+encoding+method+for+Census+2016

PART D – PRIVACY IMPACTS AND APP COMPLIANCE

This chapter analyses how MADIP impacts privacy using the APPs as a framework. The analysis assesses the compliance of MADIP with the APPs and also examines privacy impacts beyond APP requirements through the consideration of community expectations and privacy best practice. The analysis is based on:

- The OAIC *Australian Privacy Principles guidelines* and *Guide to undertaking privacy impact assessments*;
- The 2018 PIA;
- Privacy advice from Maddocks and engagement with the MADIP Board; and
- The stakeholder consultations undertaken by the ABS to inform this PIA Update.

The analysis against the APPs will be complemented by recommendations to improve APP compliance and suggested steps to improve privacy best practice. The analysis against the APPs includes consideration of the new types of data planned for linkage to MADIP. This chapter also includes a section on additional privacy considerations for the new types of data planned for linkage to MADIP.

4.1 APP 1 – Open and transparent management of personal information

APP 1 requires that an APP entity takes reasonable steps to ensure it complies with the APPs and has a process for dealing with complaints about its compliance with the APPs. APP 1 also requires that an APP entity has a clear, up to date, and available privacy policy. Further information on the requirements of APP 1 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Partially compliant with APP 1 with *Further measures possible*. Two recommendations for compliance were made, which have both been accepted and completed by the ABS.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	S1. The ABS should update online materials to include more information on: • Non-enduring linkages that have occurred. • Some summary information about the datasets that have been linked to MADIP. • MADIP's governance framework, including a description of the decision and assessment process (including risk assessments) for new linkages and research projects using MADIP. S2. The ABS should continue to obtain advice from a range of sources (including data custodians, whole-of-government building trust initiatives, and stakeholders such as those consulted for this PIA Update) to develop content for online materials for MADIP.

APP 1 AND MADIP

This PIA Update is being conducted to assess compliance with the APPs. The MADIP PIA forms an important part of the well-established framework of governance arrangements and protections for MADIP that ensure that privacy is protected and information is secure.

The ABS published the MADIP Privacy Policy on the ABS website in June 2018, which outlines how personal information is treated as part of MADIP. Key information that can be found in the MADIP Privacy Policy includes:

- The authority of the MADIP agencies to share personal information
- The personal information that is used in MADIP
- The management of and access to personal information
- Confidentiality (i.e. restrictions on the ABS disclosing personal information)
- Security, retention, and destruction of information
- Accessing and correcting personal information
- How to make a privacy complaint

The content of the MADIP Privacy Policy complies with the requirements of APP 1.4.

The 2016 Census Privacy Policy has also been updated to make specific reference to data integration, including a link to the data integration webpages of the ABS website:

> *We collect, hold, and use personal information in the Census for statistical and non-statistical purposes in order to carry out the functions of the ABS as legislated by the Australian Bureau of Statistics Act 1975 (ABS Act), and the Census and Statistics Act 1905 (Census and Statistics Act).*
>
> *These functions include collecting, compiling, analysing, and disseminating statistics about Australia, its population and economy. 'Compilation' includes use of Census information for data integration purposes, which involves combining data from two or more sources to create new statistics.*

The ABS also maintains a public ABS Privacy Policy, which states:

> *As part of its statistical collections, the ABS collects data from individuals, households and businesses, as well as from administrative sources.*

The ABS is in the process of reviewing the scope of the ABS Privacy Policy. The ABS will consider updating the contents of the policy to reference data integration pending the outcomes of the scope review.

A range of information about the project is available on the MADIP pages of the ABS website (www.abs.gov.au/madip). Key webpages that have been developed since the 2018 PIA to provide more detail on the project include:

- MADIP data and legislation – provides information about the data in MADIP and data that may be included in the future, the types of information shared for MADIP, and the legislation under which data are shared for MADIP.
- MADIP Research Projects – presents up-to-date information about approved research projects that use MADIP data.
- Data Integration Project Register – provides a summary description of current and historic ABS data integration projects.

Transparency and effective communication are core principles of MADIP and the MADIP Board recognises the importance of transparency and taking a privacy by design approach for MADIP. While transparency about MADIP has improved since the 2018 PIA, more work can be done to improve public communication about some details of MADIP.

The importance of transparency was also raised as a common theme in the stakeholder consultations undertaken by the ABS to support this PIA Update. Based on these discussions, some best practice suggestions for improving transparency about MADIP include publishing more information on:

- Data that has been linked to MADIP as a 'once-off'.
- MADIP's governance framework, including a description of the checking and assessment process (including risk assessments) undertaken for new linkages and research projects using MADIP data.
- Some summary information about the datasets that have been linked to MADIP.

The ABS also continues to source advice from various sources, including data custodians, whole-of-government building trust initiatives, and stakeholders such as those consulted for this PIA Update, to develop content for online materials for MADIP.

4.2 APP 2 – Anonymity and pseudonymity

APP 2 requires that individuals must have the option of not identifying themselves, or of using a pseudonym, when dealing with an APP entity in relation to a particular matter, unless an exception applies. Further information on the requirements of APP 2 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Compliant with no recommendations.	Compliant

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	Nil.

APP 2 AND MADIP

As set out in the 2018 PIA, APP 2 is not relevant to the day to day operation of MADIP. This is still true in consideration of the developments in MADIP since the 2018 PIA. Some limited anonymity is provided in relation to general browsing of the ABS website and accessing information resources related to MADIP. The ABS also provides the ability in some surveys for anonymity. Otherwise, data in MADIP is covered by the exception to anonymity and pseudonymity requirements provided under APP 2.2(b).

4.3 APP 3 – Collection of solicited personal information

APP 3 requires that any personal information collected must be reasonably necessary for one or more of the collecting APP entity's functions or activities. APP 3 imposes an additional requirement for collecting sensitive information, which states that the individual about whom the sensitive information relates must consent to the collection, unless an exception applies. Further information on the requirements of APP 3 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Partially compliant with APP 3 with *Further measures possible,* with two recommendations which have both been accepted and completed by the ABS.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	S3. The ABS should update online materials to outline the data minimisation approach, including treatment of sensitive data, for MADIP.

APP 3 AND MADIP

MADIP involves two stages of data collection:

1. The ABS collecting data (including personal information and sometimes sensitive information) as a data custodian and then using it in MADIP as the accredited Integrating Authority.

2. Other data custodians or authorised entities initially collecting data (including personal information and sometimes sensitive information), and then the ABS further collecting such information for MADIP as the accredited Integrating Authority.

The ABS' collection of personal information as a data custodian comes from data that is directly collected by the ABS, such as Census of Population and Housing data and data from some ABS household surveys.

The ABS' collection of personal information for MADIP as the accredited Integrating Authority occurs via data that is directly collected by data custodians or entities authorised by data custodians and then disclosed to the ABS. The ABS follows OAIC guidelines and takes a broad interpretation of the term 'collection' for APP 3, and applies it to data disclosed to the ABS for MADIP.

Some of the data disclosed to the ABS for MADIP does not contain personal information. The ABS still takes a cautious approach when managing this data and treats it with standards appropriate for personal information.

Decisions to collect data into MADIP are considered in line with the MADIP Strategy and policies and research priorities identified by the MADIP Board.

Data minimisation is a core principle of MADIP and applies not only to data collection, but also to the data use and access arrangements for MADIP. The ABS works with MADIP data custodians to ensure that only information that is reasonably necessary for the project is shared and used in MADIP.

Sensitive information

MADIP includes data that would be considered sensitive information where the record it is contained in would be considered personal information. This includes:

- Health information;
- Ethnicity and racial background;
- Indigenous status[7];
- Religious affiliation; and
- Sexuality[8].

Sensitive information is a sub-set of personal information and is given a higher level of protection under the APPs. For APP 3, this takes the shape of a requirement for individuals to consent to the collection of their sensitive information, unless an exception applies. The collection of sensitive information by the ABS for MADIP without consent is permitted by an exception in APP 3.4, which includes collection of sensitive information where this is authorised by an Australian law.

Authority to collect personal and sensitive information for MADIP

The ABS and MADIP data custodians take steps to ensure that all data being shared under MADIP has been collected by lawful and fair means.

The ABS is authorised to collect, compile, analyse, and publish statistics under the *Australian Bureau of Statistics Act 1975* and the *Census and Statistics Act 1905*. In particular, the ABS is authorised by these laws to undertake surveys and the Census, to collect information from other government entities, to link Census data and other information, to produce statistics for analysis, and to publish statistical outputs. The ABS also complies with the *Privacy Act 1988 (Cth)* and handles personal information in accordance with the Australian Privacy Principles. Data collected by the ABS for MADIP is protected by the secrecy provisions of the *Census and Statistics Act 1905*.

Data disclosed to the ABS by data custodians for MADIP is allowed based on a combination of notices to consumers (see the discussion of notices under APP 5) and exceptions in APP 3, complemented by further provisions in the legislation that governs MADIP data custodians. The overall result is that the sharing of personal information and sensitive information in MADIP is legal, in that it complies with the relevant provisions in the *Privacy Act 1988 (Cth)*, ABS legislation and MADIP data custodian legislation.

[7] Indigenous status is not a category of sensitive information in the *Privacy Act 1988 (Cth)* but is considered a type of information about ethnicity and racial background.
[8] Information on sexuality is not collected in MADIP, but it could be inferred through other information.

Online materials clarify what data are in MADIP, the types of information shared for MADIP, the legislative basis or other authority for data sharing, as well as the fact that data minimisation occurs during both data sharing and access. The latter point is also clarified in governance materials used to guide data sharing, use, and access in MADIP.

Management of sensitive information in MADIP

The 2018 PIA recognised that MADIP has a role to play in ensuring that the collection and sharing of sensitive data in MADIP is minimised to help reduce the overall risk profile of MADIP and align the project with community expectations. The 2018 PIA recommended a review of the sensitive data in MADIP, which the ABS conducted in August 2018.

The review confirmed that while current practices meet legal obligations, some changes could be made to strengthen approaches in line with privacy best practice. The review also made the following recommendations which the ABS has accepted and is actively implementing:

- Minimise data sharing so that only data that are necessary for the purposes of the project are shared and used in MADIP.
- Categorised or derived indicators for sensitive data items are used where this is feasible and unless sensitive data items in their original form are required for statistical or analytical purposes.
- MADIP project proposals require justification for requesting sensitive data items (including level of detail requested).
- Review the retention of sensitive information where there is no compelling business case for retention, or by agreement between ABS and the relevant data custodian.

4.4 APP 4 – Dealing with unsolicited personal information

APP 4 requires that where an APP entity receives unsolicited personal information, it must determine whether it would have been permitted to collect the information under APP 3. If so, APPs 5 to 13 will apply to that information. If the information could not have been collected under APP 3, and the information is not contained in a Commonwealth record, the APP entity must destroy or de-identify that information as soon as practicable, but only if it is lawful and reasonable to do so. Further information on the requirements of APP 4 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Compliant with no recommendations.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	S4. Before data are disclosed to the ABS for MADIP, the ABS should: • Provide information and assistance to data custodians to aid them in checking data for unsolicited personal information; and • Work with data custodians to decide what reasonable steps, if any, will be taken to reduce the risk of unsolicited information being disclosed to the ABS for MADIP. S5. MADIP data sharing documentation should be updated to highlight the issue of unsolicited information and to note the ABS will work with data custodians to take reasonable steps to reduce the risk of sharing unsolicited personal information with the ABS for MADIP. S6. The ABS should update online materials to provide more information about the management of unsolicited personal information in MADIP.

APP 4 AND MADIP

The data collected by the ABS for MADIP includes personal and sensitive information as outlined in APP 3. There is a risk that administrative datasets will include personal data items that were not requested, or unexpected information in a requested data item. Data sharing agreements are signed off prior to any file transfer and specify the data items to be shared, and the steps taken by the ABS if unsolicited personal information is received. This helps mitigate the risk of a notifiable data breach by lowering the chance that unsolicited personal information will be shared.

Since the 2018 PIA, the ABS has received unsolicited personal information for MADIP which has most commonly been detected in free text fields. If unsolicited personal information is received, the ABS has policies and procedures in place to securely manage the data in accordance with APP 4 and also undertake appropriate communication with stakeholders.

Staff in the Librarian and Assembler roles in the ABS Data Linkage Centre (DLC) and ABS Data Integration Assembly section (DIA) are responsible for receiving, storing, and processing data shared to the ABS for MADIP. Staff in these roles undertake routine checks on all data received for MADIP. Examples of unsolicited personal information checked for by the DLC and DIA include (but are not limited to):

- Phone numbers;
- Identification numbers (e.g. health, tax, social security, education, trustee, prison, or military identifiers)
- Email addresses;
- Indications that a person is deceased;
- Indications that a person is subject to a court order, for example a custody order, child protection order, or an apprehended violence order; and
- Names and addresses included in analytical data (noting this data are requested as linkage data but unsolicited if supplied in analytical data, in accordance with the separation principle).

All suspected or actual incidents are logged in the ABS Security Incident Reporting System (SIRS). As part of this assessment, the reporter determines whether the unsolicited information is suspected to include personal information. The ABS ICT Security Team and the ABS Privacy Team are automatically notified of the incident through SIRS.

The ABS also engages with data custodians to inform them about the incident, develop a better understanding of the incident, and agree on an approach for managing the incident. Through this process, data custodians may determine that the unsolicited information is not personal information.

If unsolicited personal information is found, APP 4 requires that the ABS determine whether the information could have been collected under APP 3 (collection of solicited personal information), and then whether the information is contained in a Commonwealth record.

While there may be situations where unsolicited personal information may be retained, the ABS takes a cautious approach for MADIP by notifying the relevant data custodian and following their directions to either:

- Securely delete the unsolicited personal information found on the dataset (either deleting values within a data item, or deleting the data item entirely); or
- Securely delete the whole dataset and request its resupply without the unsolicited personal information.

The ABS securely deletes unsolicited personal information collected for through MADIP as soon as practicable to mitigate risks of disclosing this information further.

Additionally, supply of unsolicited personal information to the ABS is unlikely to result in serious harm to the individuals to whom the data relates. This is due to the handling procedures outlined above and also because the ABS is guided by the Five Safes framework, which includes the “Safe People” element and the separation principle which means that ABS officers have undergone training to apply the separation principle and handle data safely and cannot access personal identifiers and analysis information at the same time. Staff accessing MADIP data in Librarian, Linker and Assembler roles must be Australian citizens, not be contractors, hold and maintain a BASELINE security clearance, and have signed the:

- Undertaking of Fidelity and Secrecy under the *Census and Statistics Act 1905 (Cth)*; and
- Official Secrecy and Legal Obligations Acknowledgement.

Because of this, the supply of unsolicited personal information is unlikely to result in a notifiable data breach as it would not meet the harm threshold. However, in the event that an alleged incident involving MADIP data requires escalation, ABS governance processes require it to be reported immediately to relevant data custodians, ABS ICT Security, the ABS Privacy Team, and MADIP senior executive staff. The ABS DLC is equipped with an Incident Response Manual, which broadly follows the steps outlined in the OIAC Data Breach Preparation and Response guide.

4.5 APP 5 – Notification of the collection of personal information

APP 5 requires that where an APP entity collects personal information about an individual, it must take reasonable steps to notify the individual, or otherwise ensure the individual is aware of certain matters. Further information on the requirements of APP 5 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Action required, with two recommendations for compliance, both have been accepted with one completed and one being finalised.	*Partially compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
R1. The ABS should continue to update its APP 5 collection notices, used when the ABS is collecting personal information as a data custodian, to make it clearer to individuals that their personal information may be used for data integration. R2. In relation to collection of personal information by the ABS as the accredited Integrating Authority for MADIP, the ABS should: • Advocate with entities responsible for collection notices to enhance transparency about their disclosure of personal information to the ABS for MADIP by taking reasonable steps to update notices or otherwise make individuals aware of data use. • Continue to increase transparency about the collection and use of data, including personal information, for MADIP in online materials.	S7. Data custodian agency delegates should confirm that, for their disclosure of data to the ABS for MADIP, APP 5 notification requirements have been met and the APP 6 authority for disclosure (and use by the ABS for MADIP) is identified; data custodian delegates may wish to consult with their agency Privacy Officers or otherwise authorised officers for advice on these matters. S8. MADIP data sharing documentation should be updated to provide information that confirms how APP 5 notification requirements are met for collection of data by the ABS for MADIP.

APP 5 AND MADIP

As noted in APP 3, most of the data in MADIP is initially collected by entities other than the ABS and then disclosed to the ABS for MADIP. MADIP also includes some data that is directly collected by the ABS, such as Census data, and then used in MADIP. This secondary collection and secondary use of personal information creates some complexity with assessing compliance with, and analysing privacy risks associated with, APP 5.

The compliance of MADIP with APP 5 and associated privacy risks are assessed below for these different types of data collection for MADIP:

- The collection of personal information by the ABS as a data custodian that is then used in MADIP

- The initial collection of personal information by entities that disclose that information to the ABS for MADIP
- The collection of personal information by the ABS as the accredited Integrating Authority for MADIP from entities that initially collect it

Data in MADIP that is collected by the ABS as a data custodian (Census and ABS survey data)

The ABS provides collection notices for the statistical collections it undertakes to help meet its obligations under APP 5. Prior to the 2018 PIA, most ABS collections were supported with a notice that contained some information about the particular collection and also broad terms such as that the information would be used for statistical and research purposes. The ABS recognises that it is not likely that individuals would understand their personal information is being used for MADIP based on these broad terms.

The ABS has made significant developments in improving its collection notices since the publication of the 2018 PIA to support the potential use of personal information from these collections in MADIP. The ABS is reviewing and subsequently updating its collection notices to improve transparency about the potential use of information collected for data integration. The ABS conducts testing with interviewers and survey respondents to develop collection notices. The ABS Privacy Team is consulted on the development of ABS collection notices to ensure APP 5 requirements are met.

Updates to privacy collection notices for ABS household surveys are being rolled out. The Survey of Disability Ageing and Carers 2018 was the first out in the field to include an updated collection notice. The notice included a reference to data integration and a link to the data integration section of the ABS website – where the MADIP webpages are housed.

The collection notice wording for the 2021 Census is being developed and the ABS intends to include references that Census information may be used for data integration.

To ensure compliance with APP 5, the ABS should continue to update its APP 5 collection notices to make it clearer to individuals that their personal information may be used for data integration.

As an additional step to ensure that individuals are aware of how the ABS is using personal information it collects in MADIP, the ABS published the MADIP Privacy Policy in June 2018. This policy provides information on the following topics in relation to MADIP:

- Authority to share personal information
- What personal information is used in the MADIP
- Management of and access to personal information
- Confidentiality
- Security, retention and destruction of information
- Accessing and correcting personal information
- Making a privacy complaint
- Availability of the Policy
- Links to find more information and contact details

Data that is initially collected by data custodians other than the ABS

Data that is disclosed to the ABS for MADIP may be collected by the data custodian entity for that data or by another entity authorised by the data custodian.

APP entities that collect personal information are responsible for compliance with APP 5 including determining reasonable steps they need to take to notify individuals or otherwise ensure they are aware that the entity may disclose personal information to the ABS for MADIP.

Some entities, such as state and territory government agencies, that collect personal information and share it to the ABS for MADIP may not be APP entities. Such entities are responsible with compliance with relevant jurisdictional privacy legislation, as well as entity and data specific legislation.

The disclosure of personal information to the ABS for MADIP is further considered in APPs 3 and 6.

The collection of data by the ABS from data custodians as the accredited Integrating Authority for MADIP

As noted in APP 3, the collection of personal information by the ABS from data custodians for MADIP is still considered a 'collection' in terms of the *Privacy Act 1988 (Cth)*. This means that APP 5 applies to the ABS for this secondary collection of information.

The ABS cannot update the collection notices of other APP entities that have data in MADIP. It is also not reasonable for the ABS to directly notify, such as through a letter, each individual that provides information for these collections. Instead, the ABS relies on the collection notices of the entities that share data with the ABS for MADIP, other steps these entities may take to notify individuals, and other steps the ABS takes to build awareness of the collection and use of personal information in MADIP.

MADIP does not currently have a process for documenting the APP 5 collection notices for data shared to the ABS, and considering the scale of data sharing in MADIP, it may not be practical to document the content of notices for all data disclosed to the ABS for MADIP.

However, the sharing of data to the ABS by data custodians for MADIP only occurs where authorised delegates from each party have approved the sharing of the data. This is managed through the data sharing arrangements and governance documentation for MADIP.

Generally, when an authorised entity and the ABS are approving the sharing of data with the ABS for MADIP, both parties are acknowledging that the disclosure of data by the data custodian, and the associated collection and use by the ABS for MADIP, are permitted under relevant legislation.

Generally, the data disclosed to the ABS by authorised entities for MADIP is collected with an APP 5 notice that alerts data providers that information collected may be shared or used for secondary purposes. These notifications usually use broad terms, for example, they may say that 'information may be disclosed to other government agencies for statistical and research purposes'. It is not likely that data providers would understand their information is being shared for MADIP based on APP 5 notices that use broad terms such as these.

In response to a recommendation from the 2018 PIA, some data custodians for MADIP are taking steps to improve transparency about the disclosure of data to the ABS for MADIP:

- The Department of Education has updated its departmental privacy statement to include that personal information is disclosed to the ABS for MADIP.
- The Department of Human Services departmental short form privacy notice recognises that personal information may be used for research purposes. A revision to the Privacy Policy is in progress to explicitly recognise data sharing with the ABS, including for MADIP.
- The Department of Social Services has specified in its Privacy Policy that it collects data for a variety of different purposes including policy development, research, and evaluation.
- The Australian Taxation Office and the Department of Health are reviewing their privacy notices and are actively considering the implementation of amendments.

As the accredited Integrating Authority for MADIP and as an agency with extensive experience in collecting personal information, the ABS can take a leadership role by working with agencies to build transparency about MADIP. In particular, for the secondary collection of personal information the ABS undertakes for MADIP, the ABS should:

- Advocate with entities responsible for collection notices to enhance transparency about their disclosure of personal information to the ABS for MADIP by taking reasonable steps to update notices or otherwise make individuals aware; and
- Continue to increase transparency about the collection and use of data, including personal information, for MADIP in online materials.

The ABS takes the above to represent reasonable steps in the circumstances to comply with APP 5 for its collection of personal information from data custodians for MADIP as the accredited Integrating Authority.

4.6 APP 6 – Use or disclosure of personal information

APP 6 requires that an APP entity only use or disclose personal information for the particular purpose for which it was collected (known as the 'primary purpose'), or for a secondary purpose if the person has consented or if an exception applies, such as where the secondary use or disclosure is required or authorised by or under an Australian law. Further information on the requirements of APP 6 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Partially compliant with APP 6 with *Further measures possible*. It made two recommendations which have both been accepted by the ABS and are currently being finalised.	*Partially compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
R3. MADIP data sharing documentation should be updated to provide information that confirms the APP 6 authority for sharing data to the ABS for MADIP.	*S7 repeat –* *Data custodian agency delegates should confirm that, for their disclosure of data to the ABS for MADIP, APP 5 notification requirements have been met and the APP 6 authority for disclosure (and use by the ABS for MADIP) is identified; data custodian delegates may wish to consult with their agency Privacy Officers or otherwise authorised officers for advice on these matters.* S9. The ABS should update and maintain online materials that communicate to the public about the data shared in MADIP.

APP 6 AND MADIP

Each data custodian involved in MADIP, which is an APP entity, collects personal information reasonably necessary as part of its core functions, discloses that information to the ABS for MADIP based on a person's consent or as authorised by law, for its use or policy analysis, research, and statistical purposes.

MADIP data are used to support government decision making and academic research. It provides evidence to undertake research and evaluation activities, improve the efficiency and targeting of government services, and make policy decisions. Authorised researchers also use this information to undertake research and provide policy advice.

The compliance of MADIP with APP 6 needs to be assessed by considering the:

- Data in MADIP that is disclosed to the ABS by data custodians;
- Data in MADIP that is used by the ABS for a secondary purpose, as the accredited Integrating Authority, to combine data and provide access to authorised users via highly secure ABS systems, and safeguarding privacy in collaboration with its partners – ensuring that no individual person is likely to be identified; and
- Unidentified data in MADIP that is used by authorised researchers (including data that is collected by the ABS (Census and ABS survey data) and that disclosed to the ABS by other data custodians).

Only data that are reasonably necessary for the purposes of MADIP are disclosed and used. To improve the transparency about the data shared and used in MADIP in line with privacy best practice, the ABS should update and maintain online materials that communicate to the public about the data shared in MADIP.

Data in MADIP that is disclosed to the ABS by data custodians

The ABS collects data from data custodians, or entities authorised by data custodians, for MADIP. MADIP data sharing documentation requires that these data custodians or authorised entities identify:

- If any data items that will be shared contain personal and sensitive information;
- The legislative authority for sharing data with the ABS; and
- The reason for sharing data with the ABS.

The disclosure of data to the ABS for inclusion in MADIP is not currently authorised by any general law. Unless such disclosure falls within a primary purpose for collection by the data custodian, the MADIP framework requires data custodians to confirm that other specific laws provide authority so that an exception in APP 6 applies. That legislation is then complemented by further provisions in the legislation that governs data disclosed to the ABS.

The broad legal authority for sharing information in MADIP can be accessed via the ABS website on the MADIP data and legislation page. For the legislative authority for sharing particular datasets, see Appendix 5.

The information shared for MADIP includes:

- Personal information, such as name and address, is only used to link the datasets together; it is not used for analytical purposes; and
- Analytical information may be also used to link datasets, but is primarily used for analysis by authorised researchers.

The ABS as the accredited Integrating Authority for MADIP relies on a number of provisions under APP 6 for using personal information for MADIP. Direct consent from individuals for use of their information for MADIP is not required and is usually not practical, and to date has not been relied upon. However the ABS does seek to inform individuals about use of their personal information by providing information about the project online, in the MADIP Privacy Policy, and in informative materials for particular collections which feed into MADIP. Public interest is also taken into account during the data sharing and research project approval processes by the ABS and data custodians.

To ensure continued compliance with APP 6, MADIP data sharing documentation should be updated to provide information that confirms the APP 6 authority for sharing data to the ABS for MADIP, so that this information can be more consistently documented. To facilitate this, data custodian agency delegates may wish to consult with their agency Privacy Officers or other authorised officers to identify the APP 6 authority for sharing data to the ABS for MADIP.

Data in MADIP that is used by the ABS

The ABS is authorised to collect, compile, analyse, and publish statistics under the *Australian Bureau of Statistics Act 1975* and the *Census and Statistics Act 1905*. In particular, the ABS is authorised by these laws to undertake surveys and the Census, to collect information from other government entities, to link Census data and other information, to produce statistics for analysis, and to publish statistical outputs. Data collected by the ABS (including information used in MADIP) is protected by the secrecy provisions of the *Census and Statistics Act 1905*.

Personal information used in linkage (either in original form, or changed into an unrecognisable form through methods such as character substitution or hashing to protect privacy) includes name, address, date of birth, and government identifiers. In particular, names are anonymised or encoded prior to linkage in MADIP. Other demographic information which does not directly identify a person (such as country of birth) may also be used to link datasets together where necessary to ensure high quality linked data.

The linked data available for analysis does not contain names and addresses.

Personal information in MADIP is not disclosed further by the ABS.

Unidentified data in MADIP that is used by authorised researchers

Personal information (e.g. name and address) are not available in files used for analysis, which are designed and only made available in a way that ensures an individual is not likely to be identified, in line with ABS legislation.

Access to MADIP data is only provided to authorised users for approved policy analysis, research, and statistical purposes. Government officers and non-government researchers authorised by the ABS and relevant data custodians have access to data for approved projects. All users are legally obliged to use data responsibly for approved purposes, comply with the conditions of access, and maintain the confidentiality of data.

Authorised researchers are able to use MADIP data, which does not contain direct identifiers and has been de-identified, in a secure IT environment. This means that no individual is reasonably identifiable and consequently MADIP does not disclose any personal information to researchers.

All researchers and projects must be approved for access under the Five Safes Framework.

4.7 APP 7 – Direct marketing

APP 7 requires that certain classes of APP entities must not use or disclose personal information for the purpose of direct marketing unless an exception applies, such as where the individual has consented. Further information on the requirements of APP 7 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Compliant with no recommendations.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	Nil.

APP 7 AND MADIP

Direct marketing is not relevant to MADIP.

4.8 APP 8 – Cross-border disclosure of personal information

APP 8 requires that before an APP entity discloses personal information to an overseas recipient, the APP entity must take reasonable steps to ensure that the overseas recipient does not breach the APPs (other than APP 1) in relation to the information, unless an exception applies, such as the individual has given informed consent. Further information on the requirements of APP 8 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Compliant with no recommendations.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	Nil.

APP 8 AND MADIP

Cross border data transfers are not currently relevant to MADIP. Data is not transferred or accessed outside of Australia for MADIP. The ABS and its secure ICT environment used for MADIP is fully located in Australia and data custodians (and authorised entities) that disclose personal information to the ABS for MADIP are also located in Australia.

4.9 APP 9 – Adoption, use or disclosure of government related identifiers

APP 9 requires that certain classes of APP entities must not adopt, use, or disclose a government related identifier of an individual as its own identifier of the individual unless an exception applies. Further information on the requirements of APP 9 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Compliant with no recommendations.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	Nil.

APP 9 AND MADIP

APP 9 does not generally apply to government agencies apart from some prescribed commercial activities undertaken by agencies. The ABS does not undertake such activities as part of MADIP.

The ABS may collect government related identifiers for MADIP. The ABS replaces government related identifiers received for MADIP with synthetic identifiers and may use these synthetic identifiers as part of its identity linking and identity verification process for MADIP. However, these identifiers are never disclosed by the ABS to any external parties. APP 9 does not apply to the ABS for this use of identifiers.

4.10 APP 10 – Quality of personal information

APP 10 requires that an APP entity must take reasonable steps to ensure that the personal information it collects is accurate, up-to-date, and complete. Further information on the requirements of APP 10 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Compliant with no recommendations.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	Nil.

APP 10 AND MADIP

Data quality is a core objective of MADIP. The ABS has extensive systems in place for ensuring that MADIP data are of high quality. The ABS is conscious of the importance of accurate data and data linking processes. In MADIP, research outputs and analysis will be relied on by a range of third parties, including policy makers and planners.

The ABS plays an important role in verifying identity links based on identifying information, which may constitute personal information, collected by the ABS as a data custodian or disclosed to the ABS by other data custodians or authorised entities. However, data quality issues are assessed on a case-by-case basis for each disclosure to ensure data is fit for purpose. There is a continual assessment of data linking processes in MADIP to assess the accuracy of data that is collected and used for MADIP.

The MADIP Privacy Policy notes that in some cases the ABS will adjust its linkage variables to ensure data quality:

> *Personal information used in linkage (either in original form, or changed into an unrecognisable form to protect privacy) includes name, address, date of birth, and government identifiers. Other demographic information which does not*

directly identify a person (such as country of birth) may also be used to link datasets together where necessary to ensure high quality linked data.

The ABS processes and systems that ensure data accuracy represent reasonable steps for the purposes of APP 10.

4.11 APP 11 – Security of personal information

APP 11 requires that an APP entity must take reasonable steps to protect personal information it holds from misuse, interference, and loss, as well as unauthorised access, modification, or disclosure. Further information on the requirements of APP 11 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Action required, with four recommendations for compliance, which have all been accepted and completed by the ABS.	*Action required*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
R4. The ABS should commit to undertaking a 2 yearly IRAP assessment of the MADIP operating environment as part of a regular program of audits of information security in MADIP. R5. The ABS should finalise and implement the MADIP Data Retention and Destruction policy.	S10. The ABS should enhance information in online materials about the data security protections in place for MADIP.

APP 11 AND MADIP

The ABS holds a significant amount of personal information in MADIP from a large number of data sources. The personal information held by data custodians that share data to the ABS for MADIP is not considered part of MADIP until it is collected by the ABS and is out of scope of this PIA Update.

The large amount of personal information and detailed analytical information about individuals that the ABS holds for MADIP poses significant security risks. Data integration creates a higher level of privacy and security risks, which include:

- The increased detail and volume of information about individuals increases the **risk of re-identification** and of **disclosing more about an individual** than they have provided in any one information source;
- The Census data and some of the administrative data in MADIP is comprehensive and **has very high population coverage** – it is not a sample of available data; and
- The value of the data to third parties is high, and the **data could be the target of external attack** (e.g. hacking or impersonation attempts).

The ABS recognises these risks and takes data security very seriously, and has a strong data protection culture and extensive experience in keeping data secure as Australia's national statistical organisation. The ABS was also the first agency to become an accredited Integrating Authority. The ABS maintains this accreditation which recognises that it has the appropriate settings, experience and capability, and arrangements to manage high risk data integration projects that involve Commonwealth data, such as MADIP.

The ABS has a robust framework of legislative, protective security, Information and Communication Technology (ICT), and data governance controls for protecting the privacy of individuals and ensuring data security in MADIP.

The MADIP Board remains committed to ensuring the security of personal information in MADIP is a core feature of current and potential future data sharing models for MADIP, and will continue to explore the capabilities that future IT environment and infrastructure developments may provide for data security.

Legislative controls

The ABS is subject to legislation protecting the confidentiality of information, including the *Census and Statistics Act 1905* which prohibits the ABS from releasing information in a manner that is likely to enable the identification of an individual and makes it a criminal offence to breach secrecy provisions. The penalty for a breach of the secrecy provisions includes fines of up to $25,200 or imprisonment for two years, or both.

The ABS handles personal information in accordance with the *Privacy Act 1988* and the Australian Privacy Principles, and abide by the High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes.

All ABS staff and authorised researchers that have access to MADIP data are required to sign a lifelong Undertaking of Fidelity and Secrecy under the *Census and Statistics Act 1905*.

Protective security controls

Upon appointment, all ABS staff undergo security checks. Additionally, a minimum BASELINE security clearance is required for staff involved in integrating MADIP data.

The ABS undertakes Regular Protective Security risk reviews to ensure that security arrangements continue to be effective.

Information and Communication Technology controls

The ABS has strong security arrangements for all information technology systems used for the MADIP, which:

- Conform with information technology security arrangements set out in the Australian Government Information Security Manual (ISM);
- Ensure that data collection, data linkage, and data assembly activities for MADIP information is only conducted by a dedicated team in an isolated secure environment – the Data Integration Next Gen Infrastructure (NGI). This environment has no external connectivity (i.e. no email, internet, etc.) to mitigate risks of data exfiltration;
- Includes a secured internet gateway which is reviewed annually by the Australian Signals Directorate (ASD); and
- Includes an ongoing program of security audits and system accreditations, including an independent assessment by the ASD.

The ASD accreditation assesses systems and procedures against a whole of government ICT compliance framework, conducted by the ASD via the Information Security Registered Assessors Program (IRAP). The ABS was accredited by the ASD in 2018, verifying that ABS systems and procedures are compliant with the Australian Government ISM and the Protective Security Policy Framework (PSPF). Implementation of any outstanding IRAP recommendations are in progress. ASD re-accreditation is required every 2 years as part of the ongoing program of security audits.

To meet this requirement, the ABS should commit to undertaking a two yearly IRAP assessment of the MADIP operating environment as part of a regular program of audits of information security in MADIP.

The ICT environment for MADIP is the same NGI environment as outlined in the 2018 PIA. Since the 2018 PIA, the NGI has received full ASD accreditation, as mentioned above. The NGI environment is regularly audited and all activity within the NGI is logged and subject to ongoing monitoring.

Data governance controls

There are two types of information collected by the ABS or disclosed to it by other agencies for use in data integration projects:

- Personal information which could directly identify a person (e.g. name, address, date of birth) is only used to enable datasets to be linked; it is not used for analytical purposes.
- Other information (e.g. occupation, income, health services use) may be used to combine datasets, and is also used for analysis.

When undertaking data integration activities, the ABS applies the separation principle to store identifiable personal information separately from other information, and to restrict project members' access to information according to what is necessary for

their function or role. A person working on a project can only hold one role at a time. This means that personal information and analytical information cannot be accessed at the same time, and no person can ever see all information together at any point in the process.

The ABS implements the separation principle via the application of functional separation. This was assessed in 2018 as part of the ongoing internal ABS audit program, conducted by an external auditing agency. As part of this successful audit, several recommendations were made to further enhance the implementation of functional separation across the program, all of which have been implemented.

The application of the separation principle for MADIP is described in the MADIP DATA AND INFORMATION FLOWS chapter of this document.

The ABS follows international best practice by using the Five Safes Framework to provide secure access to unidentified MADIP data for authorised researcher – it encompasses:

- **Safe people** – the data will only be provided to individuals employed, contracted, or sponsored by government departments. Each organisation has an Undertaking signed by their Responsible Officer, and each researcher signs an Individual Undertaking and Declaration of Compliance agreeing to abide by ABS conditions of access and use.
- **Safe projects** – projects must be for statistical or research purposes. Projects that intend to use the Causes of Death dataset will be additionally scrutinised. Where a project involves detailed cause of death data, the ABS will assess a further "Safe People" aspect: whether the analysts may have detailed knowledge of many real-world responses to this item (e.g. if they are program administrators), to ensure there is no risk of identity disclosure. If not, the project will proceed as standard under Section 15 of the *Census and Statistics (Information Release and Access) Determination 2018*. If there are any concerns, the project may proceed via other means (e.g. via secondment, or via Section 15 but with collapsed responses).
- **Safe settings** – the data will only be accessed via the secure controlled DataLab environment.
- **Safe data** – subject to the caveats on safe people and safe projects as above, the risk of spontaneous recognition has been assessed as low. The global risk assessment is applicable to any combinations of the variables in the global asset. If additional data to be released was not part of the MADIP global assessment and contains new datasets linked with MADIP that contain new quasi-identifiers, or new time periods for identified quasi-identifiers, the director of the ABS Customised and Microdata Delivery team will assess the Safe Data risk. Where additional risks are posed, the new extract will be reviewed by the ABS Disclosure Review Committee (DRC) and a recommendation will be made to the Australian Statistician or delegate. Where additional risks are not posed, the new extract will be noted for DRC information prior to release.
- **Safe outputs** – all outputs produced by researchers are vetted and cleared by ABS officers to ensure no information is released from the DataLab in a manner that is likely to enable the identification of an individual.

The ABS recognises that unnecessary retention of data once it is no longer required presents additional security risks. The MADIP Privacy Policy includes a section on security, retention, and destruction of information. It specifies that:

> *All information in the MADIP is retained by the ABS while there is a business need to do so. Both the source data that was used to combine datasets and the data that is used for analysis need to be retained in order to maintain and update the integrated data. The need for retention is reviewed annually for the project. This is consistent with the Privacy Act 1988.*

MADIP complies with ABS data retention policies which ensure that the retention of information is managed in line with the *Census and Statistics Act 1905*, *Archives Act 1983*, and *Privacy Act 1988 (Cth)*.

The MADIP Operating Model contains a more project specific data retention and destruction policy for MADIP. This policy states that:

> *Personal and sensitive information will be securely destroyed when there is no compelling business case to retain or quarantine this information, or by agreement between the ABS and the relevant data custodian. In these circumstances, information will be securely destroyed in accordance with the Archives Act 1983 and the Privacy Act 1988 (Cth). In some circumstances, the portfolio legislation of MADIP Partners and/or data custodians may still apply.*

Data custodians can request that the ABS destroy data in MADIP that they have custodianship of. The discretion for granting a request by a party for the ABS to modify, destroy, or dispose of records of original information provided for the project rests with the Australian Statistician, subject to the requirements of subsidiary agreements, the Archives Act, and the Privacy Act.

The process for secure deletion of data in MADIP involves:

- *Identifying affected datasets and all copies held by ABS;*
- *If necessary, creating of a new version of each affected dataset without the data items to be destroyed; and*
- *Securely deleting the affected datasets held by the ABS (including all historical backups and copies)*

At the time of writing this PIA Update Report, the MADIP Operating Model was under review by the MADIP Board. Through finalising the Operating Model, the ABS should finalise and implement the MADIP Data Retention and Destruction policy.

To increase transparency about the security of personal information in MADIP in line with privacy best practice, the ABS should enhance information in online materials about the data security protections in place for MADIP.

The above framework of protections represent reasonable steps to protect personal information in MADIP and assure compliance with APP 11. The ABS will continue to investigate new methods, approaches and technologies to safeguard the data in MADIP.

4.12 APP 12 – Access to personal information

APP 12 requires that an APP entity that holds personal information about an individual must give the individual access to that information on request, unless an exception applies. Further information on the requirements of APP 12 can be on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Action required, with one recommendation for compliance, which was accepted and completed by the ABS.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	Nil.

APP 12 AND MADIP

ABS has a general exemption to access requests to MADIP data. The *Privacy Act 1988 (Cth)* allows an agency to refuse access where they are authorised under Freedom of Information (FOI) legislation to refuse access to documents. Schedule 2, part II, division 2 of the *Freedom of Information Act 1982 (Cth)* exempts the ABS from providing access to documents containing information collected under the *Census and Statistics Act 1905*. This general exemption covers MADIP as the *Census and Statistics Act 1905* is also the legislation under which the ABS collects MADIP data. Each individual data custodian (and authorised entity) that discloses data to the ABS for MADIP remains responsible for managing access requests relating to their own data holdings, and this includes the ABS in relation to the data it directly collects. Where MADIP access requests are received by the accredited Integrating Authority (the ABS), it will refer these requests to relevant data custodians.

The other MADIP data custodians may not be subject to this exception, and under APP 12 they may still be deemed to 'hold' personal information where they have shared it with another agency. That is, they may retain their access request responsibilities.

The MADIP Privacy Policy outlines this point:

> *You can apply to access or correct your information held by the agency which originally collected it, however it may not be possible for the ABS to correct or provide you with access to information that has been collected as part of the Census or which has been integrated with other datasets in the MADIP.*

It also clarifies this:

> *It is also important to note that personal information such as names and addresses are removed when combined with other datasets as part of the data integration process, which makes it unlikely the ABS would be able to locate your information in the MADIP to update or correct.*

The MADIP Privacy Policy also outlines the process for contacting the ABS to enquire about accessing personal information in MADIP.

4.13 APP 13 – Correction of personal information

APP 13 requires that an APP entity must take reasonable steps to correct personal information to ensure that, having regard to the purpose for which it is held, it is accurate, up-to-date, complete, relevant, and not misleading. Further information on the requirements of APP 13 can be found on the OAIC website here.

APP COMPLIANCE

2018 MADIP PIA	MADIP PIA Update
Compliant with no recommendations.	*Compliant*

RECOMMENDATIONS AND SUGGESTIONS

Compliance recommendations	Best practice suggestions
Nil.	Nil.

APP 13 AND MADIP

The ABS has policies and procedures in place for complaints and the correction of inaccurate data.

MADIP data custodians that are APP entities are required to provide mechanisms for dealing with corrections and complaints, detailed in their respective privacy policies. Correction of source data is the responsibility of the relevant custodian.

The MADIP Privacy Policy provides information about requests to correct personal information:

> *You can apply to access or correct your information held by the agency which originally collected it, however it may not be possible for the ABS to correct or provide you with access to information that has been collected as part of the Census or which has been integrated with other datasets in the MADIP…*
>
> *It is also important to note that personal information such as names and addresses are removed when combined with other datasets as part of the data integration process, which makes it unlikely the ABS would be able to locate your information in the MADIP to update or correct.*

The MADIP Privacy Policy also outlines the process for contacting the ABS to enquire about correcting personal information in MADIP.

4.14 Further Best Practice Suggestions

The PIA Update has considered privacy impacts beyond the APPs and makes two additional suggestions to improve privacy best practice for MADIP.

The OAIC *Guide to undertaking privacy impact assessments* provides guidance about when a PIA should be conducted. It outlines that the first step is to conduct a threshold assessment to assess whether a PIA is necessary. The ABS conducts a threshold assessment for each new dataset proposed to be included in MADIP, and based on this assessment, determines whether a PIA update is required.

To be transparent about how threshold assessments are conducted, it may be useful for the MADIP Board to consider the triggers for either updating the MADIP PIA or conducting a separate PIA.

> ***Best Practice Suggestion 11:***
> *The MADIP Board should establish and document threshold triggers for future updates of the MADIP PIA.*

The consultation sessions the ABS conducted to inform this PIA Update raised the potential value in consulting broadly about MADIP and learning from other public data consultations. For example, there are opportunities to gauge community views on privacy issues through existing government consultation mechanisms, such as those conducted by the Office of the National Data Commissioner for the proposed *Data Sharing and Release Act*.

> ***Best Practice Suggestion 12:***
> *The MADIP Board should consider leveraging government consultation mechanisms to obtain broad-based advice on privacy issues and ethics issues relevant to MADIP.*

4.15 New Data Types – Discussion

Three new types of data are proposed for addition to MADIP as outlined in Section 3.3. For each of these new types of data, compliance with the APPs in general is covered in the relevant APP analysis in this Report. Following standard MADIP practice, before being linked to MADIP, datasets for each of these new types of data will be assessed to ensure specific APP requirements (such as for APP 5 and APP 6) are met.

The MADIP DATA AND INFORMATION FLOWS chapter of this Report provides a description of these new types of data in the context of particular projects that propose to include these data in MADIP.

This section will analyse these new types of data through a general privacy lens.

ABS SURVEY DATA

Scope

This PIA Update process has considered the linkage of non-sensitive ABS survey data to MADIP. The linkage of sensitive survey data to MADIP needs to be supported by a separate PIA process. For example, the 2014-15 National Health Survey (NHS) was linked to MADIP in 2018 following the conduct of an independent PIA. (Note: Sensitive data is data that would be considered sensitive information under the *Privacy Act 1988 (Cth)* if the data included personal information.)

In line with OAIC guidelines, the ABS considers that PIA processes conducted for a particular purpose can support new activities that follow the same procedures and involve a similar type of data. For example, the independent PIA conducted for the 2014-15 NHS to MADIP linkage can cover the linkage of sensitive survey data that is similar to the 2014-15 NHS data to MADIP.

The future linkage of sensitive survey data to MADIP will be conducted within the framework set by this PIA Update, as well as other PIAs supporting the sensitive data linkage.

Context

The ABS collects a significant amount of detailed information through its household survey program. Information collected typically includes socio-demographic information such as sex, age, educational attainment, in addition to topic-specific data.

Some surveys collect sensitive data, for example, the 2018 Survey of Disability, Ageing and Carers collected detailed health related information on people with disability.

Considerations

Survey data collection is quite different in nature to administrative data collection. For administrative data collections, individuals are required to provide information for a primary administrative purpose, such as to access government services or payments, or to comply with a compulsory activity like paying tax. Disclosing administrative data for linking and research is likely to be a secondary purpose. For ABS household surveys, data is collected primarily for statistical and research purposes.

The ABS does not require consent to use survey information collected under the *Census and Statistics Act* 1905 for further statistical and research purposes. Going beyond the exception provided by this legislation, it would also not be practical for the ABS to seek consent from individuals to use their information from surveys in MADIP.

Without the requirement for consent, it is particularly important to be transparent with individuals about the purposes for which their information is collected and how it will be used. Thus, one of the primary considerations for linking survey data is APP 5 (notification), and the requirement that survey respondents should be reasonably aware of the broad purposes their personal information will be used for at the point of collection.

In the PIA for the NHS 2014-15 to MADIP linkage, Galexia noted that:

- The collection notice was accurate at or about the time of collection;
- The privacy notice referred to NHS data being "used for research and statistical purposes" and that the linkage fell within the broad category of "use"; and
- The ABS Privacy Policy notes that some disclosure of data will take place.

With these considerations, Galexia concluded there was not a compliance issue with APP 5 (notification) for linking NHS 2014-15 to MADIP.

The ABS is reviewing and updating its collection notices for household surveys. For example, the 2018 Survey of Disability, Ageing and Carers (SDAC) was the first collection to include an updated privacy notice with a specific reference to data integration.

Conclusion

The ABS considers that for surveys conducted before the ABS had an operational data integration program in place, at 1 July 2018, a reference in the privacy notice to use for research and statistical purposes would be sufficient to meet APP 5 requirements for use in MADIP. For surveys conducted since 1 July 2018, collection notices should be clear that data may be used for data integration. Based on analysis in this PIA Update and in the NHS 2014-15 PIA, ABS considers the integration of NHS 2017-18 in MADIP meets APP requirements. This is because the nature of the information is equivalent to the information collected in the 2014-15 NHS and the survey was conducted before the ABS had an operational data integration program in place.

The linkage of SDAC 2018 to MADIP also meets APP requirements. The nature of the information is broadly equivalent to the information collected in the 2014-15 NHS (namely sensitive health information) and the survey privacy notice noted that the data would be used for data integration.

This PIA Update supports the future linkage of non-sensitive survey data to MADIP if it follows the framework of procedures, recommendations and suggestions described in this report, including conducting a PIA threshold assessment.

Other ABS surveys that collect sensitive data are not in scope of this PIA Update. Future projects that seek to link this kind of survey data with MADIP will need to be covered by a separate PIA process.

DETAILED ABORIGINAL AND TORRES STRAIT ISLANDER DATA

Scope

MADIP already includes some Aboriginal and Torres Strait Islander data in the form of Indigenous identifiers from Census and administrative sources. This PIA Update process has examined the linkage of detailed Aboriginal and Torres Strait Islander data to MADIP for the Outcomes for Job Seekers project outlined in Chapter 3.4.

Future research projects that seek to use detailed Aboriginal and Torres Strait Islander data in MADIP are not covered by this PIA Update.

Context

The Outcomes for Jobseekers project is seeking to link detailed information from the National Indigenous Australians Agency (NIAA) relating to the Indigenous Employment, Community Development and Remote Jobs and Communities programs with MADIP.

Information about an individual's Indigenous status is considered sensitive information[9]. This information may be used to assist linkage and may be provided in a de-identified form in analytic datasets to authorised researchers for approved projects.

De-identified information is not personal information and is subsequently not sensitive information for the purposes of the *Privacy Act 1988 (Cth)*, as it is not reasonably re-identifiable. When providing access to researchers, the ABS ensures that no individual is reasonably identifiable from the data remaining after the de-identification process through applying the data protections of the Five Safes Framework to minimise disclosure risk. This approach is appropriate given the high risk profile of the MADIP dataset and this project.

Considerations

To build understanding of the potential privacy impacts arising from the Outcomes for Job Seekers project, the ABS undertook consultation sessions with the ABS Round Table on Aboriginal and Torres Strait Islander Statistics and the Census 2021 Remote Expert Review Panel. Those participating in the consultations expressed general interest in the project and support for the potential value of the project to Aboriginal and Torres Strait Islander people, while noting several important conditions that would need to be met before the project could go ahead. These consultation sessions were conducted in collaboration with the ABS Centre of Excellence for Aboriginal and Torres Strait Islander Statistics. Further information on the outcomes of these consultation sessions can be accessed via the ABS website.

The consultation sessions highlighted the importance of Aboriginal and Torres Strait Islander communities getting access to data and having involvement in research projects, data sovereignty, ethics approvals and cultural safety assessments, and engagement in building in Aboriginal and Torres Strait Islander communities and seeking their views.

Conclusion

The stakeholders at the consultation sessions raised the following requirements for the Outcomes for Job Seekers project:

- The research reports produced through the project must be published and accessible by communities.
- The data linked through the project should be made available to other authorised researchers not directly involved in the project. (The stakeholders consulted also noted that Aboriginal and Torres Strait Islander community researchers should be able to become authorised researchers, and not need to partner with university researchers and government to gain access to MADIP.)
- The project must be supported by an assessment and ethics approval by an Aboriginal and Torres Strait Islander ethics committee, preferably through the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS).
- Aboriginal and Torres Strait Islander communities and representatives must be consulted at key stages of the project.
- Cultural safety must be assessed as a part of the project to ensure the cultural safety of the research design and outputs produced.

[9] Section 6 of the Privacy Act 1988 http://www.austlii.edu.au/au/legis/cth/consol_act/pa1988108/s6.html>.

The Round Table also raised the issue of consent for the use of detailed Aboriginal and Torres Strait Islander data in MADIP. Some stakeholders felt that individual consent was needed to link data, while others noted that seeking consent would not be practical or possible, and a requirement for consent would mean that the project would be unable to proceed. On balance, this PIA Update considers that the Outcomes for Job Seekers project should be able to proceed without specific consent, once data custodians confirm that APP 3 collection, APP 5 notification and APP 6 disclosure requirements and are satisfied and that the requirements set out in the points above are met. Further discussions with stakeholders are needed to build understanding about consent, collection, notification and disclosure.

The ABS notes the importance of appropriate management of all data, including detailed Aboriginal and Torres Strait Islander data, to maintaining and further building community support for MADIP. While the above considerations do not necessarily relate to a specific APP, the ABS recognises the importance of pursuing best practice in management of detailed Aboriginal and Torres Strait Islander data in MADIP.

BUSINESS CHARACTERISTICS OF EMPLOYERS TO EMPLOYEE DATA

Scope

This PIA Update process has considered the linkage of selected business characteristics (such as employer size, industry and turnover) to employee records in MADIP.

The ABS and several government and research stakeholders recognise the value in more detailed, longitudinal linkage of business data to MADIP. However, this future potential linkage is not considered as part of this PIA Update and may need to be covered by other PIA processes.

Context

MADIP currently only includes person-centred data that comes from information collected from persons and households. Two 2019-20 DIPA projects are proposing to link business characteristics to information about individuals in MADIP.

Considerations

The inclusion of a small set of business characteristics in MADIP does not raise particular APP compliance issues. While this is a new type of data, it will follow standard MADIP processes and approaches to governance, data provision, linkage, and access.

In some instances business data may include personal information to the extent that the data deals with sole proprietors. However, the small set of business characteristics data that will be linked to MADIP are not personal information. This is because the data items are not connected to information that would identify a sole proprietor.

As the business data that will be linked to MADIP are not personal information, the APPs do not generally apply to the collection, use or disclosure of this data. However, the ABS recognises that adding business characteristics to MADIP increases the amount of information about individuals and has associated privacy impacts. The strict protocols the ABS has for managing person-centred data outlined in this PIA Update will be used by the ABS to manage business characteristics information in MADIP.

Stakeholders consulted through the PIA Update process were generally comfortable with linking some business information to person records in MADIP and did not raise specific concerns.

Conclusion

This PIA Update supports the linkage of selected business characteristics to MADIP without any additional requirements.

PART E – NEXT STEPS

The MADIP Board will respond to the recommendations and suggestions made in this PIA Update. This response will be published on the ABS website alongside this Report. The ABS will continue to take a privacy by design approach to the management of data in MADIP. Some of the expected next steps in privacy management for MADIP are outlined below.

IMPLEMENTATION REPORT

The MADIP Board will publish a progress report (similar to the current Implementation Report) on the ABS website within one year of this PIA Update being published to inform on progress of implementing recommendations and suggestions from the PIA Update Process.

FUTURE UPDATES

MADIP will continue to adapt and evolve to policy and research priorities, methodological and technological advancements, and other environmental changes. New types and additional years of data may be linked to MADIP in response to policy priorities and research demands. The ABS is also continuously improving data handling practices and infrastructure for MADIP to preserve privacy, ensure data security, and increase data quality and utility.

Due to MADIP's evolution, the 2019 MADIP PIA Update has a defined scope for examining potential privacy impacts of changes to the project. It is anticipated that future updates will be required to examine substantial changes to MADIP after this PIA Update. Future updates will be conducted in line with the *OAIC Guide to undertaking privacy impact assessments*.

The 2019 MADIP PIA Update has identified future changes to MADIP which are likely to necessitate an update to the MADIP PIA. These may include, but are not limited to:

- Further updates to how MADIP information is stored or accessed
- Updates to the ABS', MADIP Partner Agencies', or the broader data legislative environment
- Proposals to link further new types of data using MADIP

The ABS and MADIP data custodians do not operate only within the MADIP framework, but are subject to the broader data and privacy environment. The Government is investing $65 million over four years to reform Australia's data system. The reforms include establishing the Office of the National Data Commissioner and introducing Commonwealth data sharing legislation. The National Data Commissioner is responsible for implementing a data sharing framework to improve social and economic outcomes for Australians, while safeguarding data. The Government appointed Ms Deborah Anton as the interim National Data Commissioner on 9 August 2018. While in its early stages, the proposed reforms are likely to impact on the ABS and MADIP operating environment in the future, and are a likely prompt for a future MADIP PIA Update. Further information on the proposed reforms can be accessed via: https://www.datacommissioner.gov.au/.

Similarly, should substantial changes to privacy legislation, ABS legislation, or legislation that enables MADIP data sharing, a further update to the MADIP PIA may be required.

MORE INFORMATION

For more information on the project, and to stay up-to-date with the evolution of MADIP, visit https://www.abs.gov.au/madip.

PART F – APPENDICES

Appendix 1 – Acronyms

Acronym	Term
ABS	Australian Bureau of Statistics <www.abs.gov.au>
AIHW	Australian Institute of Health and Welfare <www.aihw.gov.au>
API	Application Programming Interface
APP	Australian Privacy Principle
ARID	Address Register ID
ASD	Australian Signals Directorate <www.asd.gov.au>
ASGS	Australian Statistical Geography Standard
ATO	Australian Taxation Office <www.ato.gov.au>
BLADE	Business Longitudinal Analysis Data Environment <www.abs.gov.au/blade>
DHHS	Victorian Department of Health and Human Services <www.dhhs.vic.gov.au>
DIA	ABS Data Integration Assembly (section)
DIPA	Data Integration Partnership for Australia <www.pmc.gov.au/public-data/data-integration-partnership-australia>
DLC	ABS Data Linkage Centre
DRC	Disclosure Review Committee
ICT	Information and Communication Technology
IRAP	Information Security Registered Assessors Program <https://www.cyber.gov.au/programs/irap>
ISM	Australian Government Information Security Manual <https://www.cyber.gov.au/ism>
MADIP	Multi-Agency Data Integration Project <www.abs.gov.au/madip>
MEDB	Medicare Enrolments Database
MOU	Memorandum of Understanding
NGI	Next Gen Infrastructure
NHS	National Health Survey
NIAA	National Indigenous Australians Agency <www.niaa.gov.au>
OAIC	Office of the Australian Information Commissioner <www.oaic.gov.au>
PIA	Privacy Impact Assessment
PIT	Personal Income Tax
PSPF	Protective Security Policy Framework <www.protectivesecurity.gov.au>
SDAC	Survey of Disability, Ageing and Carers

SIRS	ABS Security Incident Reporting System
SSRI	Social Security and Related Information

Appendix 2 – Glossary

Term	Description
Accredited Integrating Authority	An agency authorised to undertake high-risk data linkage projects involving Commonwealth data for statistical and research purposes.
Administrative data	Data maintained by governments and other entities, including data used for registrations, transactions, and record keeping, usually during the delivery of a service.
Australian Privacy Principles	Principles contained in the *Privacy Act 1988* that regulate the way we collect, store, provide access to, use, and disclose personal information.
Cloud	As per the US National Institute of Standards and Technology (NIST) definition of cloud computing.
Data custodian	The agency that collects or generates data for any purpose, and is accountable and responsible for the governance of that data.
Data minimisation	The principle of only authorising the sharing and use of data that is reasonably necessary for a permitted purpose
Direct identifier	Information which, by itself, is able to identify an individual, organisation, or other entity.
Five Safes Framework	An internationally recognised approach to managing disclosure risk – each 'safe' refers to an independent but related aspect of disclosure risk.
Functional separation	A collection of access controls and procedures to restrict and regulate access to data. Staff are allocated different roles so that no one has the ability to access the identifying details of an individual at the same time as accessing other information about that individual or business.
Interoperability	The ability of data integration systems to exchange and make use of existing linkage results.
Longitudinal dataset	Linking multiple collections of the same data over different points in time.
Memorandum of Understanding	An agreement between two parties for data to be provided and used for a specific purpose.
Microdata	Data in a unit record file that provides detailed information about people, households, businesses or other types of records.
Personal information	As defined in section 6(1) of the *Privacy Act 1988*.
Person linkage spine	Provides the central index around which person-centred linkage is managed.
Privacy Impact Assessment	A systematic assessment of a project that identifies the impact that it might have on the privacy of individuals, and sets out recommendations for managing, minimising, or eliminating that impact
Public Interest Certificate	Signed and issued by a data custodian to certify that it is deemed necessary in the public interest to release data. Specifies the data to be released, the recipient, and the purposes for which the data will be released and used.
Re-identification	The act of determining the identity of a person or organisation even though directly identifying information has been removed.
Sensitive data	Data that would be considered sensitive information under the *Privacy Act 1988 (Cth)* if the data included personal information
Sensitive information	As defined in section 6(1) of the *Privacy Act 1988*.
Separation principle	No individual can access both the identifying information used for linkage, such as name, address and date of birth, together with the analytical information which does not contain direct identifiers.
Unidentified information	Information is considered 'unidentified' when direct identifiers such as name and address are removed or altered into an unidentifiable form, and other factors (such as combination of variables) are managed to ensure a person is not reasonably identifiable. In addition, access is controlled via the Five Safes Framework.

Appendix 3 – MADIP Strategy

MADIP Vision

The Multi-Agency Data Integration Project (MADIP) provides whole-of-life insights that can improve the lives of all Australians. By bringing together a broad set of person-centred data from across government domains, MADIP facilitates the use and re-use of public data for research purposes.

MADIP Objectives

To support this vision, MADIP will:

- Maintain trust in data quality and coverage through regular updates to current MADIP datasets, high-quality infrastructure, and data integration practices that set the standard across government.
- Expand its content to include datasets that align with policy priorities and deliver tangible public benefits.
- Protect privacy through best-practice information handling, data security, and other privacy preserving processes.
- Proactively build public trust, social license, and maintain project transparency.
- Support data users to find new insights by providing safe access to a range of users, a variety of data products, and delivering a high-quality user experience.
- Extend the value of existing public data resources, through interoperability with other datasets, integration across person, business and environmental domains, and other statistical innovations.

MADIP Benefits

MADIP will deliver a range of benefits for:

- **The public**: MADIP supports decisions that will help Australians live healthier, happier, and more independent lives.
- **Data providers:** MADIP makes better use of information that has already been collected. By combining administrative datasets, survey datasets, and the Census, MADIP enhances the value of existing public data resources.
- **Government agencies**: With its unique breadth and coverage, MADIP is a powerful tool for informing government decision-making.
- **Researchers**: MADIP makes a wider range of data available to inform research by individual, organisations and academics.

MADIP Operating Principles

The MADIP asset and system will operate in accordance with the following principles:

- MADIP is a **person-centred** data asset. The MADIP asset is structured around a spine where each unit represents one natural person.
- MADIP contains high-value datasets that are **relevant** to significant policy priorities.
- MADIP data is **regularly updated** and provides **comprehensive coverage** of the Australian population.
- MADIP operates in an **ethical** manner within the boundaries of **social licence** and community expectations.
- MADIP access is bound by the constraints of the **legislation** of its data custodians, as well as the *Census and Statistics Act 1905* and the *Privacy Act 1988*.
- MADIP data is **accessible** and **useable**, supported with metadata and other explanatory material.
- MADIP governance mechanisms are **responsive** to the changing data environment and **complement** other strategic fora.
- MADIP **supports exploration and ongoing development** to enable rapid responses to emerging risks and opportunities.

Appendix 4 – Stakeholder consultation

During the development of this PIA we met with the following organisations:

- Australian Bureau of Statistics
- ABS Round Table on Aboriginal and Torres Strait Islander Statistics
- Australian Child Rights Taskforce
- Australian Council of Social Services
- Australian Institute of Health and Welfare
- Australian National University
- Australian Privacy Foundation
- Australian Taxation Office
- Census 2021 Remote Expert Review Panel
- Commonwealth Treasury
- Consumers Health Forum
- Department of Education
- Department of Health
- Department of Human Services
- Department of Industry, Innovation and Science
- Department of Prime Minister and Cabinet
- Department of Social Services
- Families Australia
- Information Governance ANZ
- Life Course Centre (ARC Centre of Excellence for Children and Families over the Life Course)
- Melbourne Institute of Applied Economic and Social Research
- Mental Health Australia
- National Health and Medical Research Council
- New South Wales Information and Privacy Commission
- NSW Data Analytics Centre
- Office of the Australian Information Commissioner
- Office of the Victorian Information Commissioner
- Open Government Forum
- Universities Australia
- UNSW Centre for Big Data Research

Appendix 5 – Datasets in MADIP

DATASET NAME	REFERENCE PERIOD	CUSTODIAN	DESCRIPTION	LEGISLATIVE (OR OTHER) AUTHORITY	VULNERABLE POPULATION [10]	SENSITIVE DATA	ENDURING	DIPA
DATASETS CURRENTLY LINKED TO MADIP ***# INDICATES NEW LINKAGE SINCE PIA CONDUCTED IN 2017-18***								
AUSTRALIAN EARLY DEVELOPMENT CENSUS (AEDC)	2009-2016, 2018	Australian Government Department of Education	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*.	Children, Aboriginal and Torres Strait Islander	Y	Y	Y
AUSTRALIAN CENSUS LONGITUDINAL DATASET #	2006-2011-2016	ABS	Brings together a five per cent sample from the 2006 Census with corresponding records from the 2011 & 2016 Censuses	*Census and Statistics Act 1905*	Children, Aboriginal and Torres Strait Islander, Disability, Aged	Y	Y	N
AUSTRALIAN APPRENTICESHIPS INCENTIVES PROGRAM (AAIP) & TRAINING CONTRACTS #	2011-2016	Australian Government Department of Education	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 publically available. Section I.E of the 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.	Aboriginal and Torres Strait Islander, Disability	Y	Y	Y

[10] Includes the following populations: Children, Aboriginal and Torres Strait Islander, Disability, Aged.

CENSUS OF POPULATION AND HOUSING (CENSUS)	2011, 2016	ABS	Demographic information such as family composition, education attainment, marital status and household income	*Census and Statistics Act 1905*	Children, Aboriginal and Torres Strait Islander, Aged, Disability	Y	Y	Y
CENTRALISED REGISTER OF MEDICAL PRACTITIONERS (PROVIDER DIRECTORY)	2011 - 2016	Australian Government Department of Health	Information about registered medical practitioners, including specialities	Data is disclosed to the ABS pursuant to public interest certificates (PICs) issues by the Minister (delegate) in accordance with section 130 of the *Health Insurance Act 1972* and section 135A of the *National Health Act 1953*	Children, Aged, Disability	Y	Y	Y
CHILD CARE MANAGEMENT SYSTEM #	2013 - 2018	Australian Government Department of Education	Administrative data covering enrolment & attendance of children aged 4-6 [inclusive] & their associated carers, including basic demographics This information is reported to the Australian Government Department of Education and used to calculate the child care fee reductions to be paid to the service	Data is disclosed to the ABS pursuant to a public interest certificate (PIC) issued in according with paragraph 168(1)(a) of the *A New Tax System (Family Assistance) (Administration) Act 1999*	Children, Aboriginal and Torres Strait Islander	Y	N	Y
DATA EXCHANGE #	2015 - 2018	Australian Government Department of Social Services	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*	Disability, Aboriginal and Torres Strait Islander	Y	N	Y
DEATHS REGISTRATION #	2005-2017	State & Territory Registrars of Births, Deaths and Marriages[11]	All States & Territories death information collected from medical Certificates of Cause of Death, Death Information	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,	Aboriginal and Torres Strait Islander	Y	Y	Y

				and Marriages (RBDMs) are authorised to collect this data by the Births, Deaths and Marriages Act 1995 (NSW) and the Births, Deaths and Marriages Registration Act 1996 (VIC), - 2003 (QLD), - 1996 (SA), - 1998 (WA), - 1999 (TAS), - 2015 (NT), - 1997 (ACT). These Acts also authorise the Registrars to provide access to information for appropriate purposes.				
HIGHER EDUCATION INFORMATION MANAGEMENT SYSTEM #	2005 - 2016	Australian Government Department of Education	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*	Children, Aboriginal and Torres Strait Islander, Disability	Y	Y	Y
HOUSEHOLD INCOME & EXPENDITURE SURVEY #	2015 - 2016	ABS	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*	Disability	Y	N	N
MEDICARE BENEFITS SCHEDULE (MBS)	2005-2016	Australian Government Department of Health	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.	Children, Disability, Aged	Y	Y	Y

MEDICARE ENROLMENTS DATABASE (MEDB)	2006-2016	Australian Government Department of Health, Department of Human Services	Information on persons enrolled with Medicare	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*.	Children, Aged	Y	Y	Y
MIGRATION DATA - CLIENT INFORMATION #	1984 - 2018	Australian Government Department of Home Affairs	Client information on Australian-born citizens, temporary & permanent migrants	*Migration Act 1958* & the *Australian Border Force Act 2015*		Y	Y	Y
MIGRATION DATA - SKILLED MIGRATION POINTS #	2005 - 2019	Australian Government Department of Home Affairs	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*		Y	Y	Y
MIGRATION DATA - TRAVELLER DATA #	2004 - 2018	Australian Government Department of Home Affairs DHA	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*		Y	Y	Y
MIGRATION DATA - VISA INFORMATION & CITIZENSHIP GRANTS #	2000 - 2018	Australian Government Department of Home Affairs	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*		Y	Y	Y
NATIONAL ASSESSMENT PROGRAM – LITERACY AND NUMERACY (NAPLAN) (QUEENSLAND) #	2010 - 2018	Queensland Department of Education	QLD NAPLAN data for state school students who were in Year 3 during 2010, 2011 & 2012	*Education General Provisions Act 2006*	Children	N	N	Y
NATIONAL HEALTH SURVEY #	2014 - 2015	ABS	Information about the health of Australians, including:	*Census and Statistics Act 1905*	Disability	Y	Y	N

			▪ prevalence of long-term health conditions ▪ health risk factors such as smoking, overweight & obesity, alcohol consumption & physical activity & ▪ demographic & socioeconomic characteristics					
NEW SOUTH WALES APPRENTICESHIPS & TRAINEESHIPS #	2011- 2016	New South Wales Department of Industry, New South Wales Education Standards Authority, TAFE New South Wales, New South Wales Department of Education	New South Wales education prelinked datasets including the National Assessment Program – Literacy and Numeracy (NAPLAN), TAFE, Vocational Education and Training (VET), Apprenticeships, & higher education data, integrated with existing 2011 Census, Personal Income tax data & Social Security & Related Information data - MADIP linkages with NSW – MEDB concordance	*Education Standards Authority Act 2013* & the *Data Sharing (Government Sector) Act 2015*	Aboriginal and Torres Strait Islander, Disability	Y	N	N
NSW CANCER REGISTRY #	1972-2015 1988-2016 1996 - 2017 (Or latest available immediately prior to data transfer)	Cancer Institute of New South Wales	New South Wales Linked Cancer dataset includes: Cancer Registry BreastScreen PapTest	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	Aboriginal and Torres Strait Islander, Aged	Y	N	N
PERSONAL INCOME TAX (PIT)	2006-2016 2010 2016	Australian Taxation Office	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	*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.		Y	Y	Y

	2010-2016		*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					
PHARMACEUTICAL BENEFITS SCHEDULE (PBS)	2006-2016	Australian Government Department of Health	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.	Children, Disability, Aged	Y	Y	Y
SOCIAL SECURITY AND RELATED INFORMATION (SSRI)	2010/11 - 2015/16	Australian Government Department of Social Services	Characteristics of recipients of government payments such as Age Pension, Newstart Allowance, and Family Tax Benefit	Data is provided pursuant to PICs issued under the following provisions which allow for the disclosure of personal and sensitive information where it is in the public interest: • S208 of the *Social Security (Administration) Act 1999*; • S168 of the *A New Tax System (Family Assistance) (Administration) Act 1999*; • S128 of the *Paid Parental Leave Act 2010*; and • S355 of the *Student Assistance Act 1973*.	Aboriginal and Torres Strait Islander, Disability, Aged	Y	Y	Y
SURVEY OF DISABILITY, AGEING & CARERS #	2018	ABS	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*	Disability, Aged	Y	Y	N
TRANSGENERATIONAL DATASET #	1987 - 2016	Australian Government Department of Social Services	Links the social assistance records of a birth cohort of young Australians to that of their parents	•S202(2C) of the *Social Security (Administration) Act 1999*; •S168(1)(a) of the *A New Tax System (Family Assistance) Administration Act 1999*;	Children, aged	Y	N	Y

				•S355(1)(a) of the *Student Assistance Act 1973*; & s128 (1)(a) of the *Paid Parental Leave Act 2010*. • Public Interest Certificate supplied by the Department of Social Services. (Signed 10 July 2019)				
VICTORIAN LINKAGE MAP #	1991 - 2018	Department of Health and Human Services (Vic)	A system of inked records that are identified as belonging to the same person across 30 different Victorian health & human services datasets	*Health Records Act 2001* (Vic) & the *Privacy & Data Protection Act 2014* (Vic)	Children, Aboriginal and Torres Strait Islander, Disability, Aged	Y	N	N
DATASETS TO BE LINKED TO MADIP IN FUTURE **** TBC*** **THESE DATASETS ARE PROPOSED FOR LINKAGE TO MADIP IN THE FUTURE. THE LEGISLATIVE (OR OTHER) AUTHORITY, ENDURING OR NON-ENDURING STATUS AND DIPA OR NON-DIPA STATUS IS STILL TO BE DETERMINED.**								
***AUSTRALIAN IMMUNISATION REGISTER**		Australian Government Department of Health	National register that records vaccinations given to people of all ages in Australia					
***BUSINESS LONGITUDINAL ANALYSIS DATA ENVIRONMENT (BLADE) – DATASET ***		ABS	Dataset that combines business tax data and information from ABS surveys.					
***CHILD CARE SUBSIDY SCHEME**		Australian Government Department of Education	Data from the system that manages payment and administration of the Child Care Subsidy, including accepting actual attendance times.					
***COMMUNITY DEVELOPMENT EMPLOYMENT PROJECTS (CDEP) - COMMUNITY DEVELOPMENT PROGRAM**		Australian Government Department of the Prime Minister and Cabinet	Information about the disadvantages job seekers experience in moving off income support, specific to employment services in remote Australia					

*** COMMUNITY DEVELOPMENT EMPLOYMENT PROJECTS (CDEP) - INDIGENOUS EMPLOYMENT PROGRAM**		Australian Government Department of the Prime Minister and Cabinet	Information about the disadvantages job seekers experience in moving off income support, specific to employment services in remote Australia					
*** COMMUNITY DEVELOPMENT EMPLOYMENT PROJECTS (CDEP) - REMOTE JOBS & COMMUNITIES PROGRAM**		Australian Government Department of the Prime Minister and Cabinet	Includes information about the disadvantages job seekers experience in moving off income support, specific to employment services in remote Australia					
CHILD PROTECTION & OUT OF HOME CARE		Victoria Department of Health and Human Services	Clients that are involved in a child protection investigation or are in Victorian State care					
***NATIONAL EARLY CHILDHOOD EDUCATION & CARE**		State & Territory Departments of Education	Statistics on children enrolled & attending preschool programs across Australia, in addition to data about service providers					
EMPLOYMENT SERVICES SYSTEM		Australian Government Department of Employment, Skills, Small and Family Business	Administrative data from Employment Services System					
NATIONAL DISABILITY INSURANCE SCHEME (NDIS)		National Disability Insurance Agency	Scheme participants living in NSW & ACT & their characteristics, care pathways & outcomes, & how they compare with other groups					
***NEW SOUTH WALES FAMILY AND**		New South Wales Department	Clients that are involved in a child protection investigation or are in NSW State care					

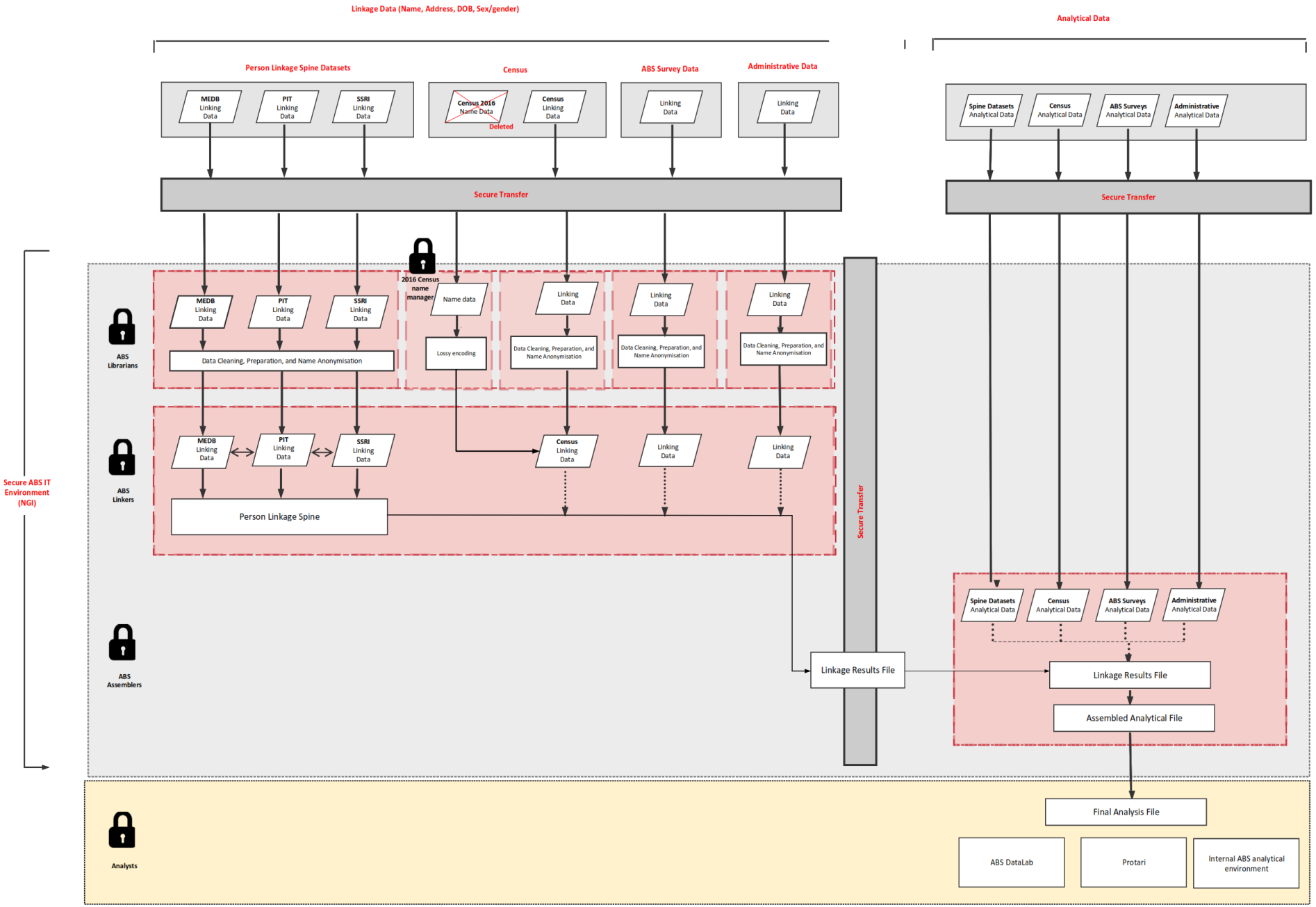
COMMUNITY SERVICES (FACS) DATA		of Family and Community Services						
NEW SOUTH WALES HIGHER SCHOOL CERTIFICATE; RECORD OF SCHOOL ACHIEVEMENT		New South Wales Education Standards Authority	Includes student & family characteristics & education characteristics					
NATIONAL ASSESSMENT PROGRAM – LITERACY AND NUMERACY (NAPLAN) (OTHER JURISDICTIONS)		Other State/ Territory Education Departments	State and Territory student level data from the National Assessment Program – Literacy and Numeracy (NAPLAN)					
NEW SOUTH WALES SCHOOL & TEACHER CHARACTERISTICS		New South Wales Education Standards Authority	School & Teacher characteristics					
NEW SOUTH WALES STUDENT OUTCOMES SURVEY		New South Wales Department of Industry	Student demographic & education characteristics					
NEW SOUTH WALES TAFE STUDENT CHARACTERISTICS, QUALIFICATIONS, ENROLMENT & ATTAINMENT		New South Wales TAFE	Student characteristics, Qualifications, enrolment & attainment data					
NEW SOUTH WALES TEACHER ACCREDITATION		New South Wales Education Standards Authority	School & Teacher characteristics					
NEW SOUTH WALES UPPER SECONDARY SCHOOL		New South Wales Department of Education	Student, family, teacher & school characteristics, Upper secondary school data (year 10,11 & 12) & Higher education data, including enrolments &					

			attendance, course data, field of education					
NEW SOUTH WALES VOCATIONAL EDUCATION AND TRAINING (VET) FUNDED PROVIDER COLLECTION		New South Wales Department of Industry	Student demographic & enrolment/education characteristics & Vocational Education and Training provider characteristics					
NEW SOUTH WALES VOCATIONAL EDUCATION AND TRAINING (VET)IN SCHOOLS		New South Wales Education Standards Authority	Student & family characteristics & education characteristics					
POST PROGRAM MONITORING SURVEYS		Australian Government Department of Employment, Skills, Small and Family Business	Survey of job seekers assisted through employment services					
TOTAL VOCATIONAL EDUCATION AND TRAINING ACTIVITY		Australian Government Department of Employment, Skills, Small and Family Business	Sourced from both the National VET Provider Collection & National VET in Schools Collection. Covers students who undertook nationally recognised Vocational Education and Training on a government funded or fee-for-service basis					

Appendix 6 – MADIP Information Flow

Next page

MADIP Data Flow



Appendix 7 – Older Australians Project Information Flow