PRIVACY IMPACT ASSESSMENT

Their Futures Matter – Integrating data from the NSW Human Services Dataset with MADIP

December 2021

Contents

PART 1 - INTRODUCTION

1.1 Background

Their Futures Matter Reform

Their Futures Matter (TFM) is a major New South Wales (NSW) government reform that aims to improve outcomes for vulnerable children, young people, and families. The reform was established in 2016 in response to a 2015 independent review of Out of Home Care. The review found that the out of home care system in NSW was *'ineffective and unsustainable, failing to improve the long-term outcomes for children or to arrest the devastating cycles of intergenerational abuse and neglect'*[1].

An integral part of the TFM reform is the 'Investment Approach', which is about *'using data and evidence to understand, prioritise and evaluate support for vulnerable children, young people and families with the highest needs, both now and into the future'*[2]. A long-term goal of the investment approach is to move away from a crisis-driven support approach towards an early intervention or prevention approach.

TFM Human Services Data Set

The Human Services Data Set (HSDS) is essential to the Investment Approach. The HSDS integrates more than 60 NSW Government datasets, from over 27 years. It includes data on child protection, housing, justice, health, education, mental health, alcohol and drug treatment, parental risk indicators and Commonwealth services such as welfare, subsidised health services and medications provided via the Medical Benefits Schedule and Pharmaceutical Benefits Scheme (MBS and PBS respectively).

The NSW Department of Communities and Justice (DCJ) commissioned Taylor Fry to analyse the HSDS data and author the 'Forecasting Future Outcomes – Stronger Communities Investment Unit 2018 Insights Report'.

In 2020, NSW DCJ approached the Australian Bureau of Statistics (ABS) to propose the Their Futures Matter project. The TFM project links a subset of the HSDS to Commonwealth data in the Multi-Agency Data Integration Project (MADIP) to produce findings that inform priority setting and resource allocation across government, with the view to improve the long-term outcomes for vulnerable children, young people and their families. This Privacy Impact Assessment (PIA) examines the privacy considerations of integrating the NSW data with MADIP.

MADIP is a partnership among Australian Government agencies to combine information on healthcare, education, government payments, personal income tax, and population demographics to create a comprehensive picture of Australia over time. An independent PIA was conducted for the MADIP asset in 2018 and updated in 2019; more information can be found in the MADIP PIA Update.

[1] Our Story | Their Futures Matter (nsw.gov.au)
[2] About the Investment Approach | Their Futures Matter (nsw.gov.au)

Need and Benefit to Link MADIP and HSDS

This project will improve understanding of vulnerable children, young people, and their families, including their government service use, and cost of these services. Coarse assumptions by NSW government suggest that around two-thirds of services provided to vulnerable people is paid for by the Commonwealth Government, including welfare, subsidised health services and medications (MBS and PBS). NSW DCJ aim to estimate more accurately the per-person cost of providing (NSW and Commonwealth) services such as welfare and primary health care to each vulnerable young person. By better understanding these costs, NSW and Commonwealth governments can invest more effectively in the right services for those who need them most.

The TFM project will create a rich dataset that enables analysts to investigate questions such as:

1. What personal and family characteristics, life events and other life variables increase the likelihood of poor future employment outcomes, welfare receipt and poor health outcomes? Understanding this helps identify potential risk factors for children to experience negative social, health, and economic outcomes into adulthood.
2. What is the interaction between welfare receipt and other key government services? For example, are past interactions with the Justice, Child Protection and Health system correlated with future welfare receipt and vice versa? This gives an understanding of compounding effects and is also useful for designing potential interventions.
3. How does the future cost of health and welfare support vary by segments of the young population? This is important in assessing potential cost savings of proposed interventions.

Adding MADIP data will enable three key improvements to the dataset:

1. Improved targeting of services for vulnerable people. Combining Commonwealth data on vulnerable people's welfare payments with NSW Government data on factors such as smoking during pregnancy, residence in social housing and incarceration will enable more appropriate service provision than is possible with NSW data alone.
2. Better understanding of the early development and health service usage pathways of children and young people, especially those who are affected by mental illness.
3. Understanding the total investment possible for a group of people. The DCJ expect to analyse the impact of welfare payments on potential investments able to be made to vulnerable people.

The TFM project agencies expect public benefits of reforms from the data platform to include:

- Improving outcomes for vulnerable children, young people, and their families, including:
    - Improved education outcomes,
    - Reduced rates of incarceration,
    - Reduced episodes of mental illness,
    - Fewer children in out-of-home-care and
    - Increased employment rates.
- Improving outcomes for the general population, such as reduced crime rates.
- Improved outcomes for this population will lower the cost of providing services to vulnerable young people and families, allowing the saved funds to be reallocated.

1.2 Purpose, scope, and approach

Purpose

The purpose of this PIA is to:

- Consider the potential privacy impacts on people whose personal information has been provided to the ABS and whose de-identified information will be made available to authorised researchers for the TFM project;
- Identify privacy risks in relation to the Australian Privacy Principles (APPs), NSW Health Privacy Principles (HPPs), and community expectations;
- Account for community attitudes towards collection and use of personal information for projects such as the TFM (see Section 1.6 Addressing community expectations); and
- Identify, assess, and outline risk mitigation strategies to manage privacy impacts.

Scope

The TFM PIA covers the one-off linkage of selected HSDS datasets to the MADIP, the data flows and processes involved, and the protections for managing personal and sensitive information. Further detail about the data and data flows is provided in Part 2 of this report. The PIA also assesses the overall compliance of the TFM project with the APPs. The TFM PIA builds on existing approved processes reviewed in the MADIP PIA Update (November 2019) and Cloud DataLab PIA (June 2020).

The NSW Government has made two Public Interest Directions (PIDs) to allow NSW agencies to disclose identified person-level data to the Data Linkage Centre (DLC) at the Centre for Health Record Linkage (CHeReL). The PIDs also permit the DLC to disclose identifiers such as names and addresses, to an Australian Government Linkage Agency (i.e. the ABS) for linking to Australian Government data. More information is available in the Public Interest Direction[3] and Health Public Interest Direction[4] made by the NSW Privacy Commissioner for the TFM Project.

Approach

The ABS followed the Office for Australian Information Commissioner's (OAIC) Guide to undertaking privacy impact assessments in completing this PIA. This included:

- Undertaking a Privacy Threshold Assessment (PTA) to determine the need for the PIA based on a shortlist of risk criteria
- Mapping information flows
- PIA analysis and compliance check
- Addressing risks
- Development of this report
- Publishing a PIA summary on the ABS website

[3] Made under section 41(1) of the *Privacy and Personal Information Protection Act 1998* (NSW).
[4] Made section 62(1) of the *Health Records and Information Privacy Act 2002* (NSW).

1.3 Compliance Summary

The PTA process found that, based on the ABS Privacy Policy, Commonwealth arrangements for data integration and Commonwealth Data Integration Risk Assessment Guidelines, the TFM project was considered high risk. This rating is partly due to the amount of sensitive information available for analysis, as outlined in Section 1.4. The TFM project's high-risk rating also comes from the complexities associated with integrating and managing a large and detailed pre-linked data asset such as the HSDS, as well as the number of agencies involved in the project (e.g. NSW DCJ, CHeReL, Taylor Fry, ABS and data custodians).

1.4 Personal and Sensitive Information

Personal Information

Definitions

The *Privacy Act 1988 (Cth)*[5] defines personal information as *"…information or an opinion about an identified individual, or an individual who is reasonably identifiable…"*[6].

The legal definition of personal information is provided by section 4 of the *Privacy and Personal Information Protection Act 1998 (NSW) Act*. Section 4 of the Act defines 'personal information' as:

"Information or an opinion (including information or an opinion forming part of a database and whether or not in a recorded form) about an individual whose identity is apparent or can be reasonably be ascertained from the information or opinion".

The *Health Records and Information Privacy Act 2002 (Part 1, section 5)* definition of personal information is similar:

"Information or an opinion (including information or an opinion forming part of a database and whether or not recorded in a material form) about an individual whose identity is apparent or can be reasonably be ascertained from the information or opinion".

While all deceased persons are out of scope of the *Privacy Act 1988 (Cth)*, the *NSW Health Records and Information Privacy Act 2002* and the *Privacy and Personal Information Protection Act 1998 (NSW)* includes those who have died in the last 30 years. For this PIA all information collected is treated as personal information, irrespective of whether the person is deceased. This is consistent with ABS practice to ensure all information collected under the *Census and Statistics Act 1905 (Cth)* is kept in a secure environment and used only for the purpose of the *Act*.

Use of Personal Information in this project

Personal information is used in the project to support analyses of vulnerable populations and identify key areas where support can be targeted to help these groups.

[5] https://www.oaic.gov.au/privacy/the-privacy-act/
[6] https://www.oaic.gov.au/privacy/guidance-and-advice/what-is-personal-information/

The NSW data is identified personal information, including name and address, to enable linking to MADIP to produce the data asset described in Section 1.1 of this PIA. The data will be handled in line with ABS security protocols as outlined in Section 2.

The MADIP PIA Update addresses the use of identifying personal information in detail, and the same will apply to the NSW TFM project:

> *"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). The ABS uses the Five Safes Framework to manage disclosure risks associated with providing access to de-identified MADIP data."*[7]

Sensitive Information

The Privacy Act defines 'sensitive data' as information or an opinion about an individual's:

- Racial or ethnic origin
- Political opinions
- Health information
- Religious affiliation
- Sexuality
- Criminal record

The NSW data and MADIP data both contain sensitive information. MADIP contains information on racial or ethnic origin, and health information (in the form of medical records). The NSW data contains information on racial and ethnic origin, health information, interactions with the criminal justice system, and interaction with child protection systems.

The ABS applies the following principles for managing sensitive data:

- Only collecting and sharing the minimum amount of personal information required for the purposes of the project (the minimisation principle);
- Using categorised or derived indicators for sensitive data items where feasible, unless sensitive data items in their original form are required for statistical or analytical purposes;
- Project proposals requiring specific justification for requesting sensitive data items; and

[7] MADIP PIA Update, pages 17-18.

Section 2.2 provides more detail on the flow of information during the linkage process.

As a project with a research focus that includes vulnerable children, the NSW TFM project must comply with the ABS Child Safety and Wellbeing policy. The policy establishes child safe practices across the ABS and details the responsibilities and obligations of ABS staff and contractors when dealing with children or young people. It is a requirement of the Commonwealth Child Safe Framework (CCSF) and reflects the National Principles for Child Safe Organisations.

1.5 Legislation and Consultation

As the Accredited Integrating Authority for MADIP, the ABS will be responsible for linking the NSW data with MADIP. The NSW data is collected by the ABS under section 10(3) of the *Census and Statistics Act 1905*. The *Census and Statistics Act 1905* provides the ABS with the legislative authority to collect data on a range of matters. The *Census and Statistics Act 1905* requires the ABS to compile such information, including through linkage, and to publish and disseminate results of such compilations and analyses, while maintaining the confidentiality of the information provided.

As a Commonwealth organisation, the ABS is bound by the *Privacy Act 1988 (Cth),* including the APPs. Compliance with the APPs is assessed in Part 3 of this PIA.

The Minister for Family and Community Services has consulted with the NSW Privacy Commissioner and obtained two PIDs for this linkage. This grants exemption to the NSW data custodians from compliance with one or more NSW Information Privacy Principles (IPPs) or HPPs under section 41(3) of the *NSW Privacy and Personal Information Protection Act 1998 (PPIP)* and section 62(3) of the *NSW Health Records and Information Privacy Act 2002 (HRIP),* which states that the Privacy Commissioner can only make a public interest direction where satisfied that the public interest in requiring the agency comply with the IPPs or HPPs is outweighed by the public interest in making the direction.

Further information on the PIDs is available here.

1.6 Addressing community expectations

While the collection, use or disclosure of personal information, including sensitive information, may be authorised by legislation, this does not necessarily mean it meets community expectations. A key privacy consideration is the right for individuals to be aware of how their personal information is being used. A PIA should therefore consider community attitudes and expectations regarding the project's privacy implications and risks. This supports building and maintaining public trust.

The OAIC's Australian Community Attitudes to Privacy Survey 2020 is a good indication of expectations held by those whose privacy may be impacted by the project. According to the report: '*privacy is a major concern for 70% of Australians, and almost 9 in 10 want more choice and control over their personal information*'. Misuse of information which doesn't seem relevant to the collection purpose is another primary concern. The ABS applies several security controls for the TFM project to mitigate these concerns. These controls are detailed in Part 2 and 3.

Public trust is critical to the ABS' reputation and to people's willingness to participate in ABS and government projects. Important steps in building public trust in the TFM project include:

- Transparency (such as listing the project on the Data Integration Project Register);
- Showing that public benefits from linking will outweigh potential privacy risks; and
- Explaining how privacy risks will be mitigated.

A key finding of the community consultation for the MADIP PIA Update was;

"…the importance of communicating the benefits of data integration to help the community understand MADIP and how it is used. Suggestions to achieve this included publishing more outputs and case studies, especially those demonstrating a tangible public benefit that has resulted from the use of integrated data"

Use of statistical information is central to making evidence-based policy decisions, which lead to better outcomes for the community. Linking the NSW data and MADIP data for the TFM project will enrich the information available about vulnerable children, young people, and their families, and draw additional value from previously collected data. The TFM project aims to provide public benefit through using the linked data to improve long-term outcomes for vulnerable children, young people, and their families. The Attitudes to Personal Information and Data Study Wave 2 (nsw.gov.au) showed that NSW residents were generally more comfortable with personal information being shared or used by government if it was for a specific reason or outcome, including assisting vulnerable children.

The ABS did not undertake additional community consultation for this PIA, given community attitudes into the use of personal information for statistical use by government have been adequately covered in other processes including:

- The MADIP PIA Update, which included community consultation;
- Approval of the two Public Interest Directions by the NSW Privacy Commissioner;
- Approval of the project proposal by all data custodians (including NSW Government departments), representing their community stakeholders; and
- NSW DJC's previous public opinion research regarding data privacy and sharing.

PART 2. DATA USE AND INFORMATION FLOWS

2.1 Data governance

The ABS is the Accredited Integrating Authority for MADIP and is responsible for receiving, storing, and linking data, assembling extracts of integrated data, and providing access to de-identified data to authorised researchers to analyse.

The usual MADIP governance processes and data infrastructure will be used to ensure data security at all stages of the TFM project. More information about these mechanisms is available in Part C of the MADIP PIA Update.

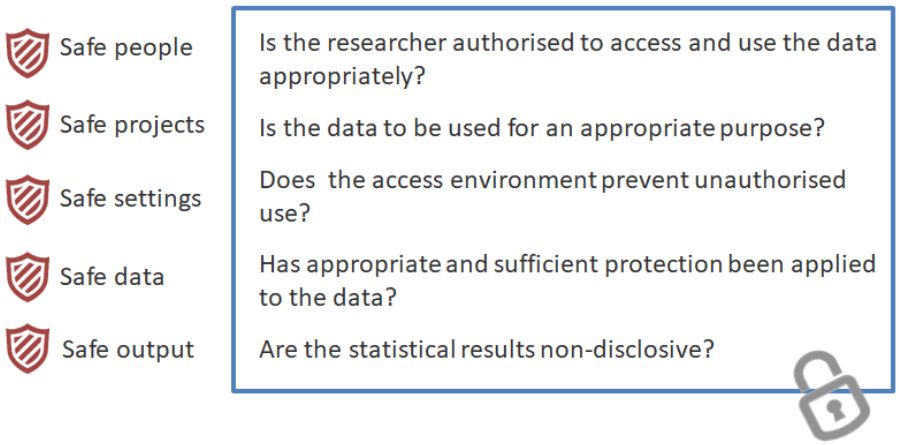
The NSW HSDS data will first be linked together by the CHeReL. Then, a subset of the pre-linked NSW HSDS data will then be requested from CHeReL using a letter of exchange and acquired by the ABS under the *Census and Statistics Act 1905.*

The NSW datasets will be supplied with a synthetic person ID to create a linked spine for the NSW data. The data will be transferred to the ABS via secure electronic means (see Section 2.2 below).

The NSW data will be linked to the MADIP spine as a one-off and will not become part of the enduring MADIP asset. All data integration activity will be performed in line with ABS functional separation principles (see Section 1.3) and conducted in the Secure ABS Data Integration Environment.

ABS' use of personal information is governed by the Five Safes Framework, as outlined on page 42 of the MADIP PIA Update.

Figure 1 Five Safes Framework

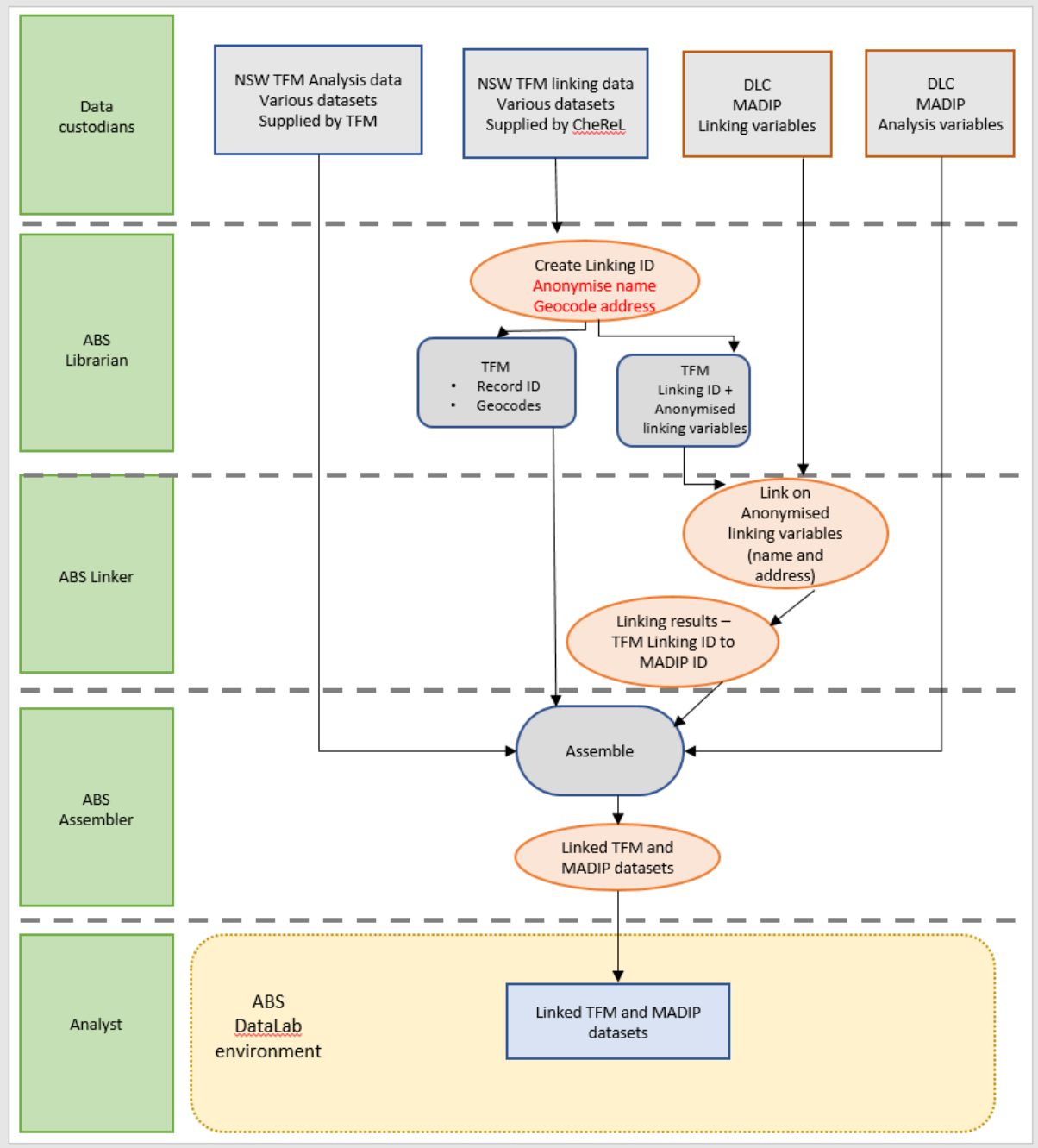


2.2 Information flows

The proposed data flows for this project, including collection, preparation, linkage, and assembly (see *Figure 2* below) follow all current standard procedures for MADIP data integration projects. Section 3.5 of the MADIP PIA Update describes in detail how the information and data in MADIP flows through the ABS environment. Practices to keep data safe include the Five Safes model for access, separation principle, and data minimisation throughout the process. This project will adhere with standard MADIP practices by ensuring:

a) All data handling will be done within the Secure Data Integration Environment (SDIE), including acquisition processes delivering newly acquired datasets directly to the SDIE.
b) Functional Separation will be followed, with only those within specific functional roles allowed to handle the data at specific designated points in the process, as per the data flow in Attachment 2.
c) Outputs will be vetted by the ABS DataLab team as per usual DataLab arrangements.
d) Data will be destroyed at the completion of the project unless prior consent is obtained from the project lead and data custodians.

Figure 2 Data flow diagram



2.3 Retention of information

The ABS recognises that retaining data when it is no longer needed presents security risks. NSW data will only be retained as agreed by the project leads and data custodians, and in compliance with obligations under the *Archives Act 1983 (Cth)*. This agreement is outlined in the Memorandum of Understanding (MOU) for the TFM project. The overarching data retention principles and arrangements set out in the MADIP PIA Update and MADIP Privacy Policy will also apply. Specifically, the MADIP Privacy Policy states that:

"*In accordance with the Australian Government records management regime, personal information in the MADIP is destroyed or deleted when no longer required. For more information, see the Administrative Functions Disposal Authority and ABS records authorities (issued in 2001 and 2007). 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.*

The TFM project does not require personal identifiers to be retained longer than the time taken to build the linked MADIP-HSDS data asset.

PART 3. AUSTRALIAN PRIVACY PRINCIPLES

APP1 – Open and transparent management of personal information

APP 1 requires that an entity manages personal information in an open and transparent way, including having a clear, up to date privacy policy that is publicly available. APP 1 also requires that an APP entity takes reasonable steps to implement practices, procedures and systems that ensure it complies with the APPs.

Compliant

The key policies for ABS management of personal information - ABS Privacy Policy for Statistical Information and the MADIP Privacy Policy - are compliant with APP 1.

The ABS MADIP Privacy Policy outlines how personal information that is collected by the ABS or shared with the ABS by data custodian agencies for MADIP (or MADIP projects) is used, and how people can access or correct their personal information. Information is disclosed to the ABS for MADIP as authorised by law for policy analysis, research and statistical purposes, consistent with the *Privacy Act 1988*. The ABS conducts MADIP in accordance with the *Australian Bureau of Statistics Act 1975* and the *Census and Statistics Act 1905*.

The *ABS Privacy Policy for Statistical Information* describes how the ABS handles personal information that is collected for producing official statistics. This includes information in datasets like the NSW HSDS that will be linked with MADIP for the TFM project. The Policy outlines:

- the kinds of personal information collected and held by the ABS
- how we collect data and keep personal information safe
- how personal information is used
- accessing and correcting personal information
- our legislative responsibilities
- how privacy complaints and enquiries can be raised and managed.

To promote transparency, all data integration projects using MADIP are summarised on the ABS' Data Integration Project Register.

APP2 – Anonymity and pseudonymity

APP 2 requires that APP entities give individuals the option of not identifying themselves, or of using a pseudonym. Limited exceptions apply.

Compliant

As outlined in the MADIP PIA Update, data used in MADIP is compliant with APP 2 due to exceptions to anonymity and pseudonymity requirements where:

- *"the APP entity is required or authorised under Australian law to do so, or*
- *if it is impractical for the APP identity to deal with individuals who have not identified themselves or who have used a pseudonym."*

It is not practicable for the ABS to deal with individuals who have not identified themselves or who have used a pseudonym as the ABS requires identified information to perform data linkage.

To protect against the risk of a data breach or disclosure risk, all names are anonymised, and addresses are geocoded. Only selected ABS staff can access identifying information, and it is kept separate from analytical information, in keeping with the functional separation principles described in Section 1.3.

APP3 – Collection of solicited personal information

APP3 covers the collection of both personal and sensitive information, the means of collections, and solicited personal information.

Compliant

Personal information other than sensitive information

APP 3.1 and 3.2 requires that any personal information collected by agency or organisation must be reasonably necessary for one of more of the collection APP entity's functions or activities.

As required by APP 3.1, all data collected by ABS for this project is reasonably necessary for the ABS to perform its function as an Accredited Integrating Authority. Personal information, such as name and address, is needed to link between the NSW data and the MADIP spine and therefore carry out the TFM project. The collected data are covered in the prescribed matters described in Section 3 of the *Census and Statistics Regulation 2016*. This includes information on education, social and welfare

services (personal but not sensitive) and crime (personal and sensitive) which are relevant to the TFM project.

Sensitive information

APP 3.3 specifies that sensitive information about an individual must not be collected by an APP entity unless:

a. The individual consents to the collection of information and:
    i. The information is reasonably necessary for, or directly related to, on or more of the agency's functions or activities; or
    ii. The information is reasonably necessary for one of more of the organisation's functions or activities; or
b. Subclause 3.4 applies in relation to the information

APP 3.4 covers examples where sensitive information may be collected. Subclause 3.4(a) outlines that sensitive information may be collected if the collection is required, authorised by, or falls under an Australian law. The MADIP PIA Update addresses disclosure of data to the ABS:

"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."[8]

The ABS manages sensitive data with the processes required by law and best practice.

Researchers must adequately justify their need to access sensitive information. Where access is granted, the ABS ensures that no individual is reasonably identifiable from the data. This approach is compatible with the 'cautious' approach recommended by the OAIC.

The MADIP PIA Update outlines how data in MADIP may be disclosed to the ABS by data custodians:

> *"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."*[9]

The legislative arrangements outlined in the MADIP PIA Update apply to the sharing and use of NSW data covered in this PIA. Additionally, exemptions to the *Privacy and Personal Information Protection Act 1998* and the *Health Records Information Protection Act 2002* have been granted by the NSW Privacy Commissioner, through the issuance of the PIDs. The exemptions mean that the TFM project

[8] MADIP PIA Update, page 30.
[9] MADIP PIA Update, Page 36.

team may share Tier Two HSDS (as defined in the PIDs) with Commonwealth Government agencies via the Data Linkage Centre with the aim of matching HSDS with Australian Government data.

A data minimisation principle is applied to all data collected by ABS for MADIP and related data integration projects so that only data that is reasonably necessary for the project is collected.

Means of collection

APP 3.5 requires that personal information may only be collected by lawful and fair means.

APP 3.6 specifies that personal information about an individual must be collected from the individual unless an exception applies:

a) the individual consents to the collection of the information from someone other than the individual; or
b) the entity is required or authorised by or under an Australian law to collect the information from someone other than the individual; or
c) it is unreasonable or impracticable for the entity to collect personal information only from the individual.

Compliant

The ABS complies with APP 3.5 because it is authorised to collect, compile, analyse, and publish statistics under the *Australian Bureau of Statistics Act 1975* and the *Census and Statistics Act 1905*. The ABS is also authorised by legislation and as an Accredited Integrating Authority to conduct projects that involve linking personal and sensitive data for statistical or research purposes.

The ABS complies with APP 3.6 because, for this project, it would be unreasonable and impracticable for the ABS to collect this data from the individuals themselves.

The NSW data has been acquired from NSW government agencies by the NSW DCJ under Section 8 (1) of the *Privacy and Personal Information Protection Act 1998 (*NSW), which requires that:

> **8 Collection of personal information for lawful purposes**
>
> *(1) A public sector agency must not collect personal information unless—*
>
> *(a) the information is collected for a lawful purpose that is directly related to a function or activity of the agency, and*
>
> *(b) the collection of the information is reasonably necessary for that purpose.*
>
> *(2) A public sector agency must not collect personal information by any unlawful means.*

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 information under APP 3. If APP 4.3

applies to the personal information, then the entity must destroy the information or ensure that it is de-identified as soon as possible, but only if it is lawful and reasonable to do so. If subclause 4.3 does not apply in relation to the personal information, APPs 5 to 13 will apply to that information.

The ABS is compliant with APP 4 regarding unsolicited personal information.

Compliant

The ABS checks all data received for unsolicited information. Unsolicited data can be in two forms:

1. Variables supplied that were not requested
2. Data within approved variables that is not of the anticipated type (e.g. phone numbers in an address field).

Data checks are made in a secure access-controlled environment. Where unsolicited data is detected, the dataset is quarantined, and an assessment is conducted to determine the nature and extent of the issue. The data custodian is advised and provided with two options:

1. ABS securely deletes the original dataset and the provider resupplies a corrected version of the data
2. The provider agrees to ABS deleting the identified unsolicited data and continuing with the use of the sanitised file.

This project will follow the standard MADIP procedures for unsolicited data, which are described in full, in the MADIP PIA Update.

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, which are outlined in APP 5.2.

Compliant

Most data used in MADIP and datasets that are linked to MADIP, including those in NSW TFM, are collected by data custodians and then disclosed to the ABS. The MADIP PIA Update includes relevant recommendations that 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 responsible steps to update notices or otherwise make individuals aware of data use' and

'continue to increase transparency about the collection and use of data, including personal information, for MADIP in online materials'.

In line with this advice, it is reasonable for the ABS to be transparent about the collection of personal information for this project by listing the project and datasets used on the Data Integration Project Register, Australia (cat. No. 1900.0). The legislation under which each of the MADIP data custodians

are authorised to collect and share personal information with the ABS for MADIP is described in the MADIP Data and Legislation page on the ABS website.

The collection and use of personal information for NSW TFM is covered by the two PIDs commissioned by the NSW Government. These PIDS have been signed by the NSW Attorney-General and NSW Minister for Health. The Information and Privacy Commission NSW has published the *Their Futures Matter PID (PPIP Act)* and *Their Futures Matter PID (HRIP Act)* on their website.

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 (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.

Compliant

Data Collected by the ABS

Personal information collected by the ABS from CHeReL, for the purposes of the NSW TFM project is collected under subsection 10(3) of the *Census and Statistics Act 1905* and the data will be used for the primary purpose it was collected. Specifically, the primary purpose is to link the NSW data with MADIP to create a limited access analytical dataset.

The ABS' Child Safety and Wellbeing policy establishes child safe practices across the ABS and details the responsibilities and obligations of ABS staff and contractors when dealing with children or young people. Personal information about children collected and held by the ABS for the NSW TFM project complies with this policy.

Data collected by NSW DCJ and supplied to CHeReL

The TFM project has released a privacy notice which relates to the collection, use, and disclosure of personal and health information. The notice states that the primary purpose of collecting data is to "deliver services and to meet our legal responsibilities". According to the PIDs, the TFM project may use personal information previously collected by an agency for a secondary purpose.

CHeReL can share the provided NSW data with the ABS under Section 27B Exemptions relating to research, *NSW Privacy and Personal Information Protection Act 1998*. This Section allows disclosure of personal information for the compilation or analysis of statistics in the public interest. The sharing is also allowed under the *Health Privacy Principle 10(f),* where information may be disclosed if "the use of the information for the secondary purpose is reasonably necessary for research, or the compilation of statistics, in the public interest".

The ABS will not disclose personal information from the NSW TFM project to anyone outside the ABS.

Access to the de-identified integrated NSW data will be provided to authorised researchers for the approved TFM project, or other research projects approved by ABS, NSW DCJ and MADIP data custodians, in accordance with standard MADIP procedures.

APP 7 – direct marketing

APP 7 requires that organisations must not use or disclose personal information for the purpose of direct marketing unless an exception applies, such as where the individual has consented.

Not applicable / Compliant

The ABS does not use or disclose personal information for direct marketing purposes.

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.

Not applicable / Compliant

Cross-border data transfers are not relevant to this project, as no data are accessed or transferred out of Australia.

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 person's government related identifier as its own identifier of the individual unless an exception applies.

Compliant

APP 9 does not generally apply to government agencies apart from some prescribed commercial activities. The ABS does not undertake commercial activities for MADIP or projects which use MADIP data; APP 9 does not apply to the ABS for this use of identifiers.

The ABS may collect government related identifiers for data linkage projects. However, these IDs are replaced with synthetic IDs as part of the linking process and original IDs are never disclosed to any external parties.

APP 10 – quality of personal information

An entity must take reasonable steps to ensure the personal information it collects is accurate, up to date, and complete.

Compliant

The ABS has extensive systems in place to ensure that personal information used in MADIP and all data linkage is of high quality.

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 continuously assess the accuracy of data linking processes and the quality of data that is provided for MADIP and the MADIP spine.

Research analysis and outputs are carried out by researchers and policy makers. While ABS officers review outputs for disclosure risk, the quality of the NSW TFM analysis is reliant on the researchers who are carrying out the analyses.

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. It must also take reasonable steps to ensure personal information is destroyed or deidentified once it is no longer needed.

Compliant

As outlined in the MADIP PIA Update:

'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 *Census and Statistics Act 1905 (Cth)* 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 security provisions.

All personal information is handled in accordance with the *Privacy Act 1988* and the *Australian Privacy Principles* and abides by the *High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes.*

All ABS staff that have access to MADIP data are required to sign a lifelong Undertaking of Fidelity and Secrecy under the *Census and Statistics Act 1905*. All authorised researchers with access to MADIP data are required to sign a legally binding undertaking that outlines a range of conditions of use, including the requirement to maintain the confidentiality of the information collected under the *Census and Statistics Act 1905*.

The ABS adheres to strong security protocols, such as functional separation, storage of data in a secure environment, and implementation of the Five Safes Framework. This project complies with the ABS data retention policy which ensures 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 PIA Update provides more information about the security of personal information in MADIP.

For NSW TFM, data is provided by CHeReL on behalf of NSW DCJ and loaded into the ABS IT environment through secure mechanisms.

The ABS and NSW DCJ have agreed to specific reporting metrics on applications, technology, staff, and third parties involved in delivering the TFM project for DCJ. These metrics are outlined in Appendix F.

Protective Security and Records Management controls

All personal information collected by the ABS is protected in accordance with the Australian Government Protective Security Policy Framework and with the Australian Government records management regime. When no longer required, personal information is destroyed or deleted according to the National Archives of Australia's Administrative Functions Disposal Authority and our records authorities (2001/00000540 and 2007/00105946).

Information and Communication Technology controls

The ABS has strong security arrangements for all information technology systems used for MADIP. Key features include:

- Arrangements which conform with Information technology security arrangements within the Australian Government Information Security Manual (ISM)
- Data integration for MADIP is carried out by a dedicated team in an isolated secure environment with no external connectivity
- A Secure Internet gateway which is independently reviewed annually by the Australian Signals Directorate (ASD); and
- An ongoing program of security audits and systems accreditations.

For more information, see the MADIP PIA Update and the Cloud DataLab PIA.

Data Governance Controls

Further protections are applied to MADIP data by use of the functional separation principle and the Five Safes Framework (described in Section 1.3).

APP 12 – access to personal information; APP 13 – correction of 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 exemption applies.

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, is accurate.

Compliant

The MADIP PIA Update and the MADIP Privacy Policy provide information on the ABS policies and procedures in place for complaints and the correction of inaccurate data. While a person can apply to access or correct information held by the agency who collected it, it may not always be possible

for the ABS to provide this access or make the corrections. This is because the relevant information may have been destroyed or personal identifiers deleted from the statistical information, which is done as soon as possible. In accordance with this policy, the ABS is not able to correct the personal information received from CHeReL for this project.

The ABS has policies and procedures in place for complaints and the correction of inaccurate data collected by the ABS under the *Census and Statistics Act 1905*. This includes data disclosed to the ABS by other data custodians for use in MADIP.

The ABS website includes current advice for accessing and correcting personal data collected under the *Census and Statistics Act 1905*:

- The *Freedom of Information Act 1982* (Schedule 2, Part II, Division 2) exempts the ABS from providing access to documents containing information collected under the *Census and Statistics Act 1905*.
- For personal information originally collected by other data custodians and shared with the ABS for data integration, each individual data custodian (and authorised entity) remains responsible for managing access requests relating to their own data holdings. They are also required to provide mechanisms for dealing with corrections and complaints, usually detailed in their respective privacy policies.

PART 4. CONCLUSION

This PIA has considered the privacy risks associated with the NSW TFM data integration project.

In undertaking this PIA, the ABS has considered the collection of the NSW data and its use for data integration purposes. Two PIDs have been issued by the NSW Privacy Commissioner where public interest is seen as outweighing compliance with NSW IPPs and HPPs.

Based on the assessment against the Australian Privacy Principles (Section 3), ABS considers that standard MADIP procedures for handling data for the purpose of integration (see Mitigation of Privacy Risks for the NSW TFRM Data Integration Project below) provide sufficient protections for the NSW TFM data integration project.

Mitigation of Privacy Risks for the NSW TFM Data Integration Project

Initial privacy risk	Description of risk	Consideration/Mitigation	Residual privacy risk
Sensitivity of data from providers **High**	The data available for analysis includes sensitive information such as medical records, indigenous status, social security information, and criminal records. There is a foreseeable risk of serious harm to individuals in the case that identification occurs either during the linkage process or during analysis or there is a breach of data security systems. In the unlikely event of a data breach, a significant amount of personal or sensitive information about a person could be involved as the project uses a wide range of information about individuals across different datasets and longitudinally.	The ABS applies the Separation Principle, Five Safes Framework, and robust data security measures to manage and protect data. Privacy management includes applying the data minimisation principle during sharing and access. Robust information security practices (independently certified as consistent with the Australian Government's Information Security Manual) are used to protect data and monitor access. Direct identifiers are altered into an unidentifiable form and/or removed during linkage to protect privacy. Only a small number of ABS officers can access direct identifiers in the MADIP (e.g. in librarian and linker roles); all are subject to lifelong secrecy obligations under the *Census and Statistics Act 1905.* Aggregate outputs of analysis (e.g. tables) undergo confidentiality checks by the ABS prior to release to prevent information being disseminated in a manner likely to enable a person to be identified.	**Medium**
Data provider consent **High**	The integrated dataset includes large amounts of administrative data generated by interactions with government services. Individuals described in the data have not explicitly consented to their information being used in this project. Public statements are available (including in collection forms and privacy policies) about the use of NSW and Commonwealth data holdings.	The NSW Privacy Commissioner issued two Public Interest Directions (PIDs) which allowed for the creation of the NSW Human Services Data Set (the Their Futures Matter PID (PIP Act) and Their Futures Matter PID (HRIP Act)), which will be linked to MADIP for the TFM project. The Directions govern the extent to which TFM and participating agencies may depart from the IPPs and HPPs for the purposes of the project. They contain a clear and defined approved purpose for which data can be disclosed, collected and used. Under the *Privacy Act 1988* personal information may be used for the purpose(s) it was collected for and for related purposes where the person	**Low**

		consents or as authorised by law. The MADIP uses data collected for research and statistical purposes for such purposes and is authorised by legislation applying to the partner agencies. Direct consent for use of information in MADIP is therefore not required or sought.	
Access to personal information in the project **Low**	Only a small number of ABS officers can access direct identifiers in the MADIP (e.g. in librarian and linker roles); all are subject to lifelong secrecy obligations under the *Census and Statistics Act 1905.* Access to the de-identified integrated NSW data will be provided to authorised researchers for the approved TFM project in accordance with standard MADIP procedures.	Access to the output file will be managed according to the Five Safes framework and the protections afforded by the *Census and Statistics Act, 1905.* The separation principle is applied so no individual will access both the linking and analytical files. Access to the analytical files will be provided to a limited number of vetted NSW government and Taylor Fry analysts via the DataLab.	**Low**
Sharing and use of personal information in the project **Low**	Personal information collected by the ABS from CHeReL, for the purposes of the NSW TFM project is collected under subsection 10(3) of the *Census and Statistics Act 1905* and the data will be used for the primary purpose it was collected. Specifically, the primary purpose is to link the NSW data with MADIP to create a limited access analytical dataset. The NSW Government commissioned two Public Interest Disclosures (PIDs) under which the TFM project may use personal information previously collected by an agency for a secondary purpose.	CHeReL can share the provided NSW data with the ABS under Section 27B Exemptions relating to research, *NSW Privacy and Personal Information Protection Act 1998*. This Section allows disclosure of personal information for the compilation or analysis of statistics in the public interest. Data sharing is also allowed under the NSW *Health Privacy Principle 10(f),* where information may be disclosed if "the use of the information for the secondary purpose is reasonably necessary for research, or the compilation of statistics, in the public interest". The TFM project has been assessed by the ABS and MADIP data custodian agencies to be in the public interest as part of the initial approvals process.	
Data quality (data is inaccurate, irrelevant, out of date etc.) **Medium risk**	Entities must take reasonable steps to ensure the personal information they collect is accurate, up to date, and complete.	The ABS has extensive systems in place to ensure that personal information used in MADIP and all data linkage is of high quality. The MADIP Privacy Policy notes that in some cases the ABS will adjust its linkage variables to ensure data quality: The ABS continuously assess the accuracy of data linking processes and the quality of data that is provided for MADIP and the MADIP spine.	**Low**

		The quality of the NSW data before delivery to the ABS is the responsibility of the CHeReL who links these data to form the HSDS asset, and the NSW government agencies contributing their data to the project.	

PART 5. APPENDICES

Appendix A – Selected HSDS Datasets, linked to MADIP

Data Provider	*Dataset included in HSDS data to be integrated with MADIP*
Department of Justice (NSW Bureau of Crime Statistics and Research (BOSCAR)	1. Finalised Charges 2. Custody episodes
NSW Department of Communities and Justice (NSW DCJ)	1. Persons file (social housing (and waitlist) and private rental assistance) 2. Social Housing Sensitive 3. Tenancy file 4. Private rental assistance file 5. Client District cumulative data 6. Child protection (FACS) 7. Out-of-home care data (FACS) 8. Homelessness services data (FACS) 9. Intensive Family Support (IFS) and Intensive Family Preservation (IFP) 10. Brighter Futures 11. Youth Hope
NSW Department of Customer Service/ Registry of Birth, Deaths and Marriages	1. Registry of Birth, Deaths and Marriages – Births 2. Registry of Birth, Deaths and Marriages – Deaths
NSW Department of Education (NSW DoE)	1. Student Details 2. Student School 3. Parent carer details 4. School attendance data 5. Suspensions
NSW Ministry of Health (NSW Ambulance)	1. Emergency Department Data Collection 2. Perinatal Data Collection (PDC) 3. NSW Ambulance Patient Health Care Record - Mental Health Ambulatory 4. NSW Ambulance Patient Health Care Record - Patient Health Care Record 5. OTP Authority 6. OTP dosing point 7. Admitted Patients Data Collection 8. NSW Minimum Data Set for drug and alcohol treatment
NSW Education Standards Authority	1. NAPLAN data 2. ROSA enrolments 3. ROSA results 4. Schools

NSW Legal Aid	1. Legal advice and minor assistance services 2. Extended legal assistance 3. Duty lawyer services 4. Grants of legal aid
NSW Police	1. Victim data 2. Juvenile justice data 3. Domestic Violence Safety assessment tool
Department of Finance Services and Innovation (Revenue NSW)	1. Vulnerable data project 2. Enforcement orders and work development orders
CHeReL	1. Concordance Table 2. Population Master List
Further data may be added as a second phase in future. This data includes NSW Department of Industry data and any other additional datasets from currently listed agencies.	NSW Department of Industry data which may be added as a second phase includes: 1. Student enrolment in funded training 2. NSW Smart and Skilled 3. Apprenticeships and traineeship

Appendix B – Acronyms

Acronym	Term
ABS	Australian Bureau of Statistics <www.abs.gov.au>
APP	Australian Privacy Principle
ATO	Australian Taxation Office
CHeReL	Centre for Health Record Linkage
DOMINO CAD	Data Over Multiple Individual Occurrences – Centrelink Administrative Data
DSS	Department of Social Services
HPP	Health Privacy Principle (NSW)
HSDS	Human Services dataset
IPP	Information Privacy Principle (NSW)
MADIP	Multi-Agency Data Integration Project <www.abs.gov.au/madip>
MBS	Medical Benefits Schedule
NSW	New South Wales
NSW data	A subset of Human Services dataset, provided for this project
OAIC	Office of the Australian Information Commissioner <www.oaic.gov.au>
PIA	Privacy Impact Assessment
PBS	Pharmaceutical Benefits Scheme
TFM	Their Futures Matter

Appendix C – 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.
Data custodian	The agency that collects or generates data for any purpose, and is accountable and responsible for the governance of that data.
De-identified	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). (De-identified data is different to unidentified data - see the meaning of unidentified data.)
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.
MADIP PIA Update	The ABS conducted a Privacy Impact Assessment of MADIP in late 2019. The MADIP PIA Update and MADIP Board response are published on the ABS website.
Microdata	Data in a unit record file that provides detailed information about people, households, businesses or other types of records.
Person Linkage Spine	The Person Linkage Spine is the central index around which person-centred linkage is managed.
Personal Information	As defined in section 6(1) of the *Privacy Act 1988*.
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.
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*.

Appendix D – Linking and analytical variables

Linking variables to be provided for all datasets

Variable	Group
Given names	Personal
Surname	Personal
Sex	Personal
Date of birth	Personal
Address	Personal

Analytical variables to be provided

HSDS Dataset Release Metadata_u

Appendix E – Documents consulted for PIA

Legislation

Commonwealth

Privacy Act 1988

Census and Statistics Act 1905

Australian Bureau of Statistics Act 1975

National Cancer Screening Register Act 2016

Archives Act 1983

State

Health Records and Information Privacy Act 2002 (NSW)

Privacy and Personal Information Protection Act 1998 (NSW)

Other PIAs:

- Privacy Impact Assessment: Cloud DataLab, ABS, 2020.
- Privacy Impact Update (PIA) for the Multi-Agency Data Integration Project (MADIP), MADIP, 2019.

Other documents:

Australian Community Attitudes to Privacy Survey 2020 (oaic.gov.au)

Public Interest Direction and Health Public Interest Direction

APPENDIX F: REPORTING METRICS

The following metrics are to be reported regarding applications, technology, staff and third parties involved in delivering the service to DCJ.

Metric details	Response (Yes, No, NA, %, value)	Reporting frequency	Target
ABS will attest at the end of each financial year that they have:			
• Undergone a security audit from a suitably independent entity or an internal department		24 months	Yes
• Conducted appropriate risk assessments of their key assets		Annual	Yes
• Provided a letter of certification from IRAP assessments of its core ICT platform		Annual	Yes
Percentage of ABS personnel (supporting DCJ services) whom have undergone appropriate background screening prior to employment – includes police check.		Annual	100%
Percentage of ABS personnel supporting the DCJ service under a valid non-disclosure agreement.		Annual	100%
Number of audits to review the physical security controls of data centres hosting DCJ data and services conducted in the last financial year.		Annual	At least 1
Has ABS enforced DCJ security requirements down the supply chain to its suppliers supporting DCJ services?		Annual	Yes
Percentage of logical access requests provisioned using a formal user registration and de-registration procedure.		Annual	100%
Percentage of terminated / resigned staff for whom the access was revoked by the last working day.		Annual	100%
When was the last access review performed of elevated access groups which support DCJ infrastructure and applications?		Annual	Within past 12 months
Percentage of servers covered by regular vulnerability assessment.		Annual	100%
Percentage of public facing servers covered by regular penetration tests.		24 months	100%
Percentage of deployed application developments which have undergone a formal security review.		Annual	100%
Percentage changes done in the production with a formal change management and security acceptance testing performed.		Annual	100%

When was the backup arrangement last successfully tested?		Annual	Date within past 12months.
Does ABS maintain an updated incident management procedure which identifies reporting of security incidents to DCJ within stipulated time-frames, which is based on the category of security incident?		Annual	Yes.
Number of security incidents impacting DCJ reported.	Critical – Important – Moderate – Low –	Annual	0