



Final: 20 September 2017

Privacy Impact Assessment Report

Vulnerable and Disadvantaged Children Research Project

Prepared for the Australian Bureau of Statistics

PIA Report: Vulnerable and Disadvantaged Children Research Project

1.	Executive Summary	3
2.	About this report	5
3.	Description of the Project	7
4.	Data sourcing	14
5.	Data Integration	21
6.	Availability of VDC data	23
7.	Other Privacy Considerations	24
	Schedule 1 – Detailed information flows	27
	Schedule 2 – Data Linkage Process Diagram	32
	Schedule 3 – Data Linkage Process: Linking and Analysis Variables	33
	Schedule 4 – NAPLAN Data Custodians	35
	Schedule 5 – SSRI Data covered by DSS PID	36
	Schedule 6 – CCMS Data covered by CCMS PIC	38
	Schedule 7 – Summary of the APPs	39
	Schedule 8 - Sources of information	41
	Schedule 9 – Glossary and acronyms	43



1. Executive Summary

1.1 Introduction

- (a) The Australian Bureau of Statistics (**ABS**) is undertaking a proof of concept project (**Project**) on behalf of the Australian Education Senior Officials Committee (**AESOC**), which involves linking certain historical datasets held by the ABS (including 2011 Census data), with other datasets provided by:
 - (i) the Australian Government Department of Education and Training (**AGDET**), including certain information sourced from the Department of Social Services (**DSS**); and
 - (ii) state and territory education departments and agencies (equivalent to AGDET), or associated National Assessment Program Literacy and Numeracy (**NAPLAN**) testing authorities (collectively referred to in this PIA as **NAPLAN Data Custodians**).
- (b) The Project will test the value of integrating the relevant datasets. The integrated data outputs will support research to better assess the impacts of current policies and programs, and shape future evidence-based policy decisions, relating to childhood education (including interventions through income support and other social and welfare services) for vulnerable and disadvantaged children.
- (c) The ABS has commissioned this independent privacy impact assessment (**PIA**) to:
 - (i) identify and assess privacy risks and impacts relating to the data integration aspect of the Project (which will be undertaken by the ABS as an Accredited Integrating Authority); and
 - (ii) identify options for managing, minimising or eliminating privacy risks and maximising privacy protection.

1.2 Summary of recommendations

- (a) Recommendations made in this PIA to mitigate identified privacy compliance risks are set out below.

Recommendation 1: To promote transparency and public debate about the benefits of, or any privacy concerns relating to, the Project, the ABS should ensure that:

- its communication strategy (which may be delivered through a range of activities and media channels) reaches as much of the Australian community as possible, and particularly parents with children; and
- easily accessible information is made available which covers the following matters:
 - why the Project is important;
 - what information will be collected and used;
 - the scope of the statistical and research purposes for which the integrated VDC data may be used;
 - an explanation of the legislative authority relied upon by the ABS to collect and use data (including the use of 2011 Population Census Data) for the purposes of the Project;
 - how the data linkage will occur; and
 - how information will be protected.

The ABS may wish to consult and/or coordinate with other Project stakeholders (e.g. education departments) in relation to the development and implementation of appropriate communication measures.



Recommendation 2: Further to Recommendation 1:

- the project description on the National Statistical Service (**NSS**) Data Integration Project register (or its replacement) should include information about the nature of the information that will be collected and used by the ABS; and
- the ABS website should include prominent and clear information about the Project, including a link to the NSS Data Integration Project register (or its replacement).

Recommendation 3: As part of the review of the ABS Privacy Policy, consideration should be given to including information about the collection of data from Commonwealth, State and Territory government bodies, and the use of that information in data integration projects.

2. About this report

2.1 Scope

- (a) The scope of this PIA:
 - (i) relates to the information flows identified in Schedule 1;
 - (ii) assesses whether the Project data integration proposal will ensure the ABS complies with its obligations under the *Privacy Act 1988* (Cth) (**Privacy Act**), including the Australian Privacy Principles (**APPs**); and
 - (iii) identifies and recommends options for managing, reducing or eliminating any negative privacy impacts that have been identified in the PIA, including any further community consultation that may be required in relation to the Project.
- (b) We are instructed that the following aspects of the Project are out of scope for this PIA:
 - (i) the provision of Stage 1 linkage datasets by the ABS to AGDET under paragraph 19(1)(b) of the *Census and Statistics Act 1905* (Cth) (**Census and Statistics Act**);¹
 - (ii) the subsequent use of the (de-identified) integrated VDC dataset to produce an analytical report for informing definitions of 'vulnerable' and 'disadvantage', and understanding the relationships between vulnerable and disadvantaged children's characteristics, ECEP experiences and longer-term outcomes, which will then inform consideration of future policy interventions.

2.2 Methodology

- (a) To produce this PIA report, MinterEllison has:
 - (i) applied an approach based on the *Guide to undertaking privacy impact assessments (May 2014)* (**PIA Guide**) issued by the Office of the Australian Information Commissioner (**OAIC**);
 - (ii) prepared a detailed narrative of the relevant information flows (as described in Schedule 1);
 - (iii) assessed the privacy impacts of the Project by reference to the APPs under the Privacy Act (summarised in Schedule 7);
 - (iv) relied on and considered the material listed in Schedule 8;
 - (v) prepared a draft PIA report for comment by the ABS; and
 - (vi) finalised the report based on further information and feedback received from the ABS.
- (b) We have not directly undertaken consultation with other agencies, stakeholders or interest groups.
- (c) Further, while the likelihood and seriousness of potential privacy risks have been considered in the context of proposing mitigation strategies, this PIA does not provide an analysis at the level of detail set out in AS/NZS ISO 31000 (Risk management – Principles and Guidelines) (although it would be possible and open to the ABS to undertake such an assessment).

2.3 Qualifications and assumptions

- (a) This PIA report is subject to the following qualifications and assumptions:
 - (i) as at the date of writing this report, the information flows described in Schedule 1 accurately reflect the current proposal as to how the Project will be implemented;

¹ Paragraph 19(1)(b) of the Census and Statistics Act provides that it is an offence for a person to divulge or communicate information given under that Act to another person, other than the person from whom the information was obtained.

- 
- (ii) MinterEllison has relied on the instructions given by the ABS regarding the scope of its powers under the Census and Statistics Act to collect, use and disclose personal information for the purposes of the Project; and
 - (iii) MinterEllison has not undertaken any consultation or investigations other than those set out in the Methodology at section 2.2.

2.4 Addressing community expectations

- (a) Authorising the collection, use or disclosure of personal information through the use of legislation or legislative instruments may ensure that a particular activity complies with the Privacy Act. However, that does not necessarily mean that it will meet community expectations. For example, community attitudes about what constitutes an invasion of privacy are not necessarily reflected in law. Accordingly, a PIA is more than just an APP compliance check, and should have regard to community attitudes and expectations when considering the broader privacy implications and risks of a project.²
- (b) Reliable indicators of community expectations are difficult to produce, and it is beyond the scope of this PIA to commission research on community attitudes relating to this particular Project.
- (c) For the purposes of this PIA, our assumptions about community expectations most likely to be held by persons whose privacy may be impacted by the project are based on the OAIC's *Community Attitudes to Privacy Survey 2017*. We have also considered submissions made in response to the Senate Economics Reference Committee's *Inquiry into the 2016 Census*, which address privacy concerns relating to the use of Census data.

² OAIC, *Guide to undertaking privacy impact assessments (May 2014)*, page 2.

3. Description of the Project

3.1 Overview of ABS functions

Statistics functions

- (a) The ABS is established under the *Australian Bureau of Statistics Act 1975* (Cth) (**ABS Act**). Its statutory functions include:³
- (i) collecting, compiling, analysing and disseminating statistics and related information;
 - (ii) coordinating with other official bodies (including other agencies) to avoid duplication in the collection, and maximise the utilisation, of information for statistical purposes.
- (b) The ABS also has functions under the *Census and Statistics Act*. Sections 8, 9 and 12 relevantly provide:

8 Taking of Census

- (1) *The Census shall be taken in the year 1981 and in every fifth year thereafter, and at such other times as prescribed.*
- (2) *The Census day shall be a day appointed for that purpose by proclamation.*
- (3) *For the purposes of taking the Census, the Statistician shall collect statistical information in relation to the matters prescribed for the purposes of this section.*

9 Statistical information to be collected

- (1) *The Statistician:*
 - (a) *may from time to time collect such statistical information in relation to the matters prescribed for the purposes of this section as he or she considers appropriate; and*
 - (b) *shall, if the Minister so directs by notice in writing, collect such statistical information in relation to the matters so prescribed as specified in the notice.*
- (2) *The Statistician shall collect such statistical information as is necessary for the purposes of the compilation and analysis, under section 12, of statistics of the number of people of each State as on the last day of March, June, September and December in each year, but nothing in this subsection shall be taken to limit the generality of subsection (1).*

12 Publication etc. of statistics

- (1) *The Statistician shall compile and analyse the statistical information collected under this Act and shall publish and disseminate the results of any such compilation and analysis, or abstracts of those results.*
- (2) *The results or abstracts referred to in subsection (1) shall not be published or disseminated in a manner that is likely to enable the identification of a particular person or organization.*
- (3) *The Statistician may make changes for results and abstracts published and disseminated under this section.*

³ ABS Act, section 6.



Release of information

- (1) *Notwithstanding anything in this Act (other than this section), the Minister may, by legislative instrument, make determinations providing for and in relation to the disclosure, with the approval in writing of the Statistician, of information included in a specified class of information furnished in pursuance of this Act.*
 - (2) *Without limiting the generality of subsection (1), determinations may make provision:*
 - (a) *as to the persons to whom the information may be disclosed;*
 - (b) *as to the persons, being the persons from whom the information has been obtained, whose consent is required for the disclosure of the information; and*
 - (c) *specifying terms and conditions subject to which the information may be disclosed, including, but without limiting the generality of the foregoing, terms and conditions as to the requiring of a person to whom the information is, or is to be, disclosed to give an undertaking, in writing with respect to the disclosure of the information by that person, including an undertaking not to disclose any of the information to any person.*
 - (3) *Information of a personal or domestic nature relating to a person shall not be disclosed in accordance with a determination in a manner that is likely to enable the identification of that person.*
- (c) For the purposes of section 9 of the Census and Statistics Act, Items 14 and 45 in the table under regulation 13 of the *Census and Statistics Regulation 2016* (Cth) (**Census and Statistics Regulation**) prescribe 'education' and 'social and welfare services' as matters in respect of which statistical information may be collected.
- (d) 'Statistical information' is not defined in the Census and Statistics Act, however the ABS considers this term means information that is used in the creation or compilation of statistics.⁴

ABS as an Accredited Integrating Authority

- (e) Data integration is the process of developing new datasets through the linkage of two or more existing datasets. The term 'data linkage' as used by the ABS, refers to the process of combining records between two datasets as part of a data integration process.
- (f) The *High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes* (**High Level Principles**) and the *Data Integration Involving Commonwealth Data for Statistical and Research Purposes: Governance and Institutional Arrangements* (**Governance Arrangements**) provide a framework for statistical data integration involving Commonwealth data. Data integration for all projects assessed as having a 'high systemic risk' (i.e. involving sensitive or personal data for statistical or research purposes) must be undertaken by an accredited Integrating Authority.
- (g) The ABS is an Integrating Authority that was independently accredited in 2012 by the Cross Portfolio Data Integration Oversight Board to undertake data integration activities. As part of the accreditation process, the ABS was audited in relation to its compliance with the following criteria:
 - (i) ability to ensure secure data management;
 - (ii) ability to ensure information that is likely to enable identification of individuals or organisations is not disclosed to external users;
 - (iii) availability of appropriate skills;
 - (iv) appropriate technical capability;
 - (v) lack of conflict of interest; and

⁴ MinterEllison has relied upon this interpretation for the purposes of preparing this PIA.

- (vi) culture and values that ensure protection of confidential information, and support the use of data as a strategic resource.

3.2 Project background

- (a) In December 2013, the Australian Government asked the Productivity Commission to undertake a public inquiry into future options for childcare and early childhood learning. The matters on which the Productivity Commission was asked to report included 'current and future needs for childcare in Australia, particularly given changes in work patterns, early learning needs, childcare affordability and government assistance'.
- (b) The Productivity Commission's final report on *Childcare and Early Childhood Learning* was released on 20 February 2015. The report considered the (then) state of Early Childhood Education and Care (**ECEC**) in Australia, and how to achieve future ECEC objectives in a way that better supports children's learning and development needs and the workforce participation of parents. The report made the following relevant recommendations:

Recommendation 5.2: Early intervention programs to address the development needs of children from disadvantaged backgrounds should be underpinned by research. Their impact on the development outcomes of children attending ECEC should be subject to ongoing monitoring and evaluation, including through the use of longitudinal studies.

Recommendation 17.7: The Australian Government should establish a program to link information for each child from the National ECEC Collection to information from the Child Care Management System, the Australian Early Development Census, and NAPLAN testing results to establish a longitudinal database. Where possible, this should also be linked to other key administration data sets and Censuses.

A confidentialised file should be made available for statistical, research, policy analysis and policy development purposes. The ability of researchers to access unit record information should be permitted subject to stringent privacy and data protection requirements.

The Australian Government Agency that is the custodian of the Child Care Management System should provide a publicly available extract from the database each year for interested parties at a sufficiently detailed geographic level for planning purposes.

- (c) In April 2016, the Productivity Commission was asked to provide advice on the national approach to collecting and using data for ECEC and schools, to improve Australia's educational outcomes. Key points from the Productivity Commission's final report on *National Education Evidence Base* (released on 24 May 2016) included the following:
- *Access to, and sharing of, education data would be substantially improved through reforms proposed in the Commission's draft report on Data Availability and Use.*
 - *Meanwhile, there is also scope to improve sharing of education data for research purposes by changing current administrative processes for collecting some education data.*
 - *There are gaps in existing data collections and work in train should fill many of them.*
 - *But the largest gaps in the national education evidence base relate to evidence, notably:*
 - *the evaluation of policies, programs and education practices in Australian schools and early childhood education and care*



services to identify what works best, for whom and in what circumstances

- *building an understanding of how to turn best practice into common practice, which is as important as evaluating what works best.*⁵

3.3 Project purpose

- (a) The purpose and drivers for the Project, as endorsed by AESOC, are described as follows:

Vulnerable and disadvantaged children face a unique set of challenges in their journey through education and life. These barriers often manifest in lower developmental and educational outcomes among vulnerable and disadvantaged children compared with their less disadvantaged peers. Through the Australian Early Development Census (AEDC) we understand that the characteristics of vulnerability and disadvantage have already manifested in the first year of full-time school.

However – we do not have a clear and measurable definition of vulnerable and disadvantaged and do not have a clear understanding of what exactly constitutes an effective intervention. It is currently challenging to respond to key Early Childhood Education and Care (ECEC) policy questions due to the disconnection between current data sets. This issue is highlighted further in the recently released Productivity Commission Draft Report – Building the National Education Evidence Base that highlights the need for improved data linkage.

To shape policy and better assess impacts of current policy and programs, we need to identify the characteristics of vulnerable families and children and the factors influencing those children's outcomes including family environment and early childhood education programs (ECEP) etc. Currently, the information available on these factors is fragmented.

This 'proof of concept' proposal seeks to provide a data linkage and analysis of the relationships between the characteristics of vulnerable and disadvantaged families and early education and care services they may have accessed. It will also provide the foundation for informing a measure of vulnerability and disadvantage, develop a more complete understanding of the relationships between children's outcomes, their family characteristics and environment, ECEP and identify risk or protective factors.

- (b) The Project extends the value of public data for the benefit of the Australian public as per the Productivity Commission's recommendations cited above.

3.4 Key project stages

- (a) The Project involves data linkage that combines the following six datasets to create a single integrated dataset (which is the focus of this PIA):
- (i) Australian Early Development Census (**AEDC**) 2012 (**AEDC Data**, described below in section 3.6);
 - (ii) National Early Childhood Education and Care Collection (**NECECC**), 2011 (**NECECC Data**, described below in section 3.7);
 - (iii) Census of Population and Housing 2011 (**2011 Census Data**, described below in section 3.8);
 - (iv) Social Security and Related Information (**SSRI**), 2005 – 2015 (**SSRI Data**, described below in section 3.9);
 - (v) Childcare Management System (**CCMS**) information, 2005 – 2015 (**CCMS Data**, described below in section 3.10); and

⁵ <http://www.pc.gov.au/inquiries/completed/education-evidence/report>.

- (vi) National Assessment Program Literacy and Numeracy (NAPLAN, 2014 – 2016 (**NAPLAN Data**, described below in section 3.11));

3.5 Information flows

- (a) A detailed description of the personal information flows are set out in Schedule 1. In brief summary:
 - (i) The ABS holds NECECC Data previously collected from AGDET, state and territory education departments and agencies, and the Catholic Education Office of the Archdiocese of Canberra and Goulburn (**Catholic Education Office**) for the 2011 NECECC.
 - (ii) For the purposes of the Project:
 - (A) AGDET will provide the AEDC Data, SSRI Data and CCMS Data to the ABS; and
 - (B) NAPLAN Data Custodians will provide the NAPLAN Data to the ABS.
 - (iii) The ABS Data Linkage Centre (**DLC**) will undertake the data linkage process in two stages (illustrated in the diagram at Schedule 2):
 - (A) In Stage 1, the SSRI and CCMS datasets will be merged, and the combined SSRI/CCMS Data will then be linked to the 2012 AEDC Data to create a de-identified dataset.
 - (B) In Stage 2, the SSRI/CCMS/AEDC 2012 Data will be linked to the NECECC, NAPLAN and 2011 Census data to create a final de-identified integrated dataset (**VDC Data**).

The linking variables and analysis variables that will be applied during the linking processes are listed in Schedule 3.
 - (iv) The final VDC Data will be made available to approved users (which may include users external to the ABS) via the ABS DataLab.
- (b) As part of the data linkage processes, the ABS will create:
 - (i) linkage input datasets, which contain the information necessary for linkage;
 - (ii) record concordances, which show the direct link between the random IDs in relation to an individual used in different datasets; and
 - (iii) linked analysis datasets, i.e. the final integrated dataset produced from linking input datasets, using record concordances and analytical variables.
- (c) These datasets in isolation may not necessarily identify individuals. However, as the ABS will have available resources to be able to ascertain the identity of individuals (i.e. by combining datasets, record concordances, etc), the data will be treated as personal information for Privacy Act purposes. As discussed later in this PIA report, there are a number of technical, procedural and environmental controls in place to restrict the circumstances in which datasets could be combined so as to enable an individual to be 'reasonably identifiable' from the information.

3.6 AEDC Data (2012)

- (a) The AEDC is a nationwide data collection of early childhood development in relation to children in their first year of school. The AEDC is held every three years, and is completed by teachers who provide information to AGDET⁶ in relation to the following five key areas of early childhood development:
 - (i) physical health and well-being;
 - (ii) social competence;
 - (iii) emotional maturity;

⁶ A parent or guardian of a child can request that their child be excluded from the AEDC.

- 
- (iv) language and cognitive skills; and
 - (i) communication skills and general knowledge.
- (b) The ABS has received data relating to the 2012 AEDC from AGDET under subsection 9(1) of the Census and Statistics Act, and a memorandum of understanding with AGDET which addresses the use of the AEDC Data for approved purposes, including data integration. The data includes personal information about children, such as their name, age, date of birth (**DOB**), sex, language spoken at home, and the teacher's assessment of the child's development and vulnerability (e.g. the child's ability to read complex words, etc). The data also includes the following 'sensitive information' (as defined in the Privacy Act) about children:
- (i) Indigenous status (being information about an individual's racial or ethnic origin); and
 - (ii) health information, for example whether the child has been assessed as having 'special needs' because they are blind, have asthma, are autistic or have another medical condition or disability.

3.7 NECECC Data (2011)

- (a) The NECECC is an annual collection of data relating to children's experience in early childhood education programs (**ECEP**), and has been compiled and collected by the ABS since 2010. The data is originally sourced by AGDET, state and territory education departments and agencies, and the Catholic Education Office, from parents and guardians of children and childhood education providers. The data is then collected by the ABS from these entities under paragraph 9(1)(a) of the Census and Statistics Act.
- (b) The Project will involve the integration of 2011 NECECC Data relating to children. This will include:
 - (i) personal information about children, such as their name, age, DOB, sex, place of residence, attendance at a preschool program (including the provider, hours of attendance and fees); and
 - (ii) sensitive information relating to the child's Indigenous status.

3.8 2011 Population Census Data

- (a) Information relating to the 2011 Population Census was collected by the ABS under section 8 of the Census and Statistics Act. The 2011 Population Census Data does not include names and addresses, but does include information relating to the background, family and dwelling characteristics of parents and their children, such as:
 - (i) the DOB, age and sex of parents and children;
 - (ii) parents' educational attainment, employment status and income;
 - (iii) parents' and children's Indigenous status and Religion; and
 - (iv) derived items including Household Income, equalised Household Income and Socio-Economic Indexes for Areas (**SEIFA**).
- (b) A copy of the 2011 Population Census form is available at [http://www.abs.gov.au/ausstats/abs@.nsf/lookup/2903.0main%20features162011/\\$file/SAMPLE_PRINT_VERSION_F1.pdf](http://www.abs.gov.au/ausstats/abs@.nsf/lookup/2903.0main%20features162011/$file/SAMPLE_PRINT_VERSION_F1.pdf).

3.9 SSRI Data (2005 – 2015)

- (a) DHS collects information about Centrelink customers, their partners and children for the purposes of administering income support and family assistance payments (including family tax benefit (**FTB**), child care benefit and child care rebate) under the social security law and family assistance law, on behalf of DSS. DSS has access to this information for the purposes of meeting its business and service delivery responsibilities in relation to the administration of the social security and family assistance laws.

- 
- (b) For the purposes of the Project, DSS will provide certain income support and family assistance information to AGDET, and AGDET will then provide the dataset to the ABS. The types of information included in the SSRI Data is listed in Schedule 5, and includes:
- (i) personal information about children, their parents and their parent's spouse;⁷
 - (ii) sensitive information about children, namely their disability status;
 - (iii) sensitive information about parents/carers and spouses, namely their Indigenous status.
- (c) The SSRI Data will also include a 'vulnerability indicator code' in relation to parents/carers. This is recorded by DHS (Centrelink) if evidence suggests that certain factors could directly impact on a job seeker's ability to comply with their income support payment obligations. This includes indicators relating to the health status of individuals, such as illness or injury requiring frequent treatment, cognitive or neurological impairment, or drug/alcohol dependency (amongst a number of other things).

3.10 CCMS Data (2005 – 2015)

- (a) The CCMS is a national online computer system which all approved child care services are required to use to record child enrolment and attendance information. This information is collected and used by AGDET and DHS to calculate FTB, child care benefit and child care rebates. A subset of CCMS child care data is included in the NECECC.
- (b) For the purposes of the Project, the ABS will obtain CCMS data from AGDET under subsection 9(1) of the Census and Statistics Act in relation to the children born between 1 January 2005 to 30 September 2015.⁸
- (c) The CCMS Data will include:
- (i) sensitive information about children, specifically their Indigenous status; and
 - (ii) other personal information about children, including their name, DOB, sex, current and previous address (if they had moved in the previous six months), and date of death (if relevant).

3.11 NAPLAN Data (2014 – 2016)

- (a) NAPLAN is an annual assessment for all students in Years 3, 5, 7 and 9 in relation to literacy and numeracy. Data is collected by the NAPLAN Data Custodians listed in Schedule 4 to this report under the *Australian Education Act 2013*.
- (b) For the purposes of the Project, and pursuant to subsection 9(1) of the Census and Statistics Act, the ABS will obtain NAPLAN Data relating to Year 3 students for the years 2014 – 2016. The personal information about students will include their name, address, DOB, sex, and achievements in literacy and numeracy. Indigenous status information will also be collected.

⁷ ABS advises that Centrelink Customer Reference Numbers (CRNs) are not provided by AGDET to the ABS (although they may be provided by DSS to AGDET).

⁸ The parents/carers for these cohorts of children were eligible for FTB Part A and/or B at the end of each quarter from when they were born.

4. Data sourcing

4.1 Legislative authority

- (a) A number of APPs provide exceptions to general obligations and limitations on the handling of personal information where it is 'required or authorised by or under an Australian law'. The OAIC's APP Guidelines provide the following guidance in relation to this exception:

Meaning of 'required'

B.129 An APP entity that is 'required' by an Australian law or a court/tribunal order to handle information in a particular way has a legal obligation to do so, and cannot choose to act differently. The obligation will usually be indicated by words such as 'must' or 'shall', and may be accompanied by a sanction for non-compliance.

Meaning of 'authorised'

B.130 An APP entity that is 'authorised' under an Australian law or a court/tribunal order has discretion as to whether it will handle information in a particular way. The entity is permitted to take the action but is not required to do so. The authorisation may be indicated by a word such as 'may', but may also be implied rather than expressed in the law or order.

B.131 An APP entity may be impliedly authorised by law to handle personal information in a particular way, where a law requires or authorises a function or activity, and this directly entails the information handling practice. For example, a statute that authorises an APP entity to collect personal information about an individual from a third party implicitly authorises the entity to disclose the individual's identity to the third party.

B.132 An act or practice is not 'authorised' solely because there is no law or court/tribunal order prohibiting it. Nor can an act or practice rely solely on a general or incidental authority conferred by statute upon an agency to do anything necessary or convenient for, or incidental to or consequential upon, the specific functions and powers of the agency. The reason is that the purpose of the APPs is to protect the privacy of individuals by imposing obligations on APP entities in handling personal information. A law will not authorise an exception to those requirements unless it does so by clear and direct language.

Census and Statistics Act – statistical information

- (b) As indicated above in paragraph 3.1(d), the ABS considers that the term 'statistical information', as used in the Census and Statistics Act, means information that is used in the creation or compilation of statistics. On that basis, subsection 9(1) of the Census and Statistics Act would provide the ABS with legislative authority to collect:
- (i) AEDC, SSRI and CCMS Data from AGDET; and
 - (ii) NAPLAN Data from NAPLAN Data Custodians.
- (c) The ABS advises that the NECECC Data was also previously collected by the ABS under subsection 9(1) of the Census and Statistics Act.
- (d) The 2011 Population Census Data was collected by the ABS under section 8 of the Census and Statistics Act.

Social Security (Administration) Act 1999

- (e) The SSRI Data comprises information obtained by DHS for the purposes of the social security law. Such information is 'protected information' subject to secrecy provisions under the *Social Security (Administration) Act 1999 (SSA Act)*.

- (f) Subsection 202(2C) of the SSA Act authorises a person to obtain, make a record of, use or disclose protected information if the Secretary of DSS (or his delegate) believes, on reasonable grounds, that this is reasonably necessary for the purposes of:
 - (i) research into matters of relevance to a Department that is administering any part of the social security law;
 - (ii) statistical analysis of matters of relevance to a Department that is administering any part of the social security law; or
 - (iii) policy development.
- (g) A delegate of the Secretary of DSS formed the view that it is reasonably necessary for AGDET and ABS officers involved in the Project to obtain, record, disclose and use SSRI data to link and analyse the relevant datasets, which will support research and policy development in relation to early childhood education (including interventions through income support and other social and welfare services). This is reflected in an instrument made by the delegate in 2017 under subsection 202(2C) of the SSA Act (**DSS PID**).
- (h) The effect of subsection 202(2C) of the SSA Act and the DSS PID is to provide a legislative authority:
 - (i) for AGDET to disclose SSRI Data (listed in Schedule 5) to the ABS; and
 - (ii) for the ABS to collect, record, use and disclose that information.⁹

A New Tax System (Family Assistance) (Administration) Act 1999

- (i) The CCMS Data comprises information obtained by AGDET for the purposes of the family assistance law. Such information is 'protected information' subject to secrecy provisions under the *A New Tax System (Family Assistance) (Administration) Act 1999 (FAA Act)*.
- (j) Paragraph 168(1)(a) of the FAA Act authorises the Secretary of AGDET (or her delegate) to disclose protected information if the Secretary certifies that it is in the public interest to do so.
- (k) The *Family Assistance (Public Interest Certificate Guidelines) Determination 2015 (PIC Guidelines)*, made under paragraph 169(a) of the FAA Act, sets out circumstances in which the Secretary's power to issue public interest certificates (**PICs**) may be exercised. Section 14AA of the PIC Guidelines provides that protected information may be disclosed:

... where the disclosure is necessary for the purpose of:

 - a) *research into (including evaluation or monitoring of, or reporting on); or*
 - b) *statistical analysis of; or*
 - c) *policy development in relation to,*

any programme or activity within the portfolio responsibilities of a department (other than a department administering any part of the family assistance law or the social security law) which is responsible for administering early childhood development, or pre-school education, policies or programmes.
- (l) In 2017, a delegate of the Secretary of AGDET made a PIC under paragraph 168(1)(a) of the FAA Act and section 14AA of the PIC Guidelines (**CCMS PIC**), which authorises AGDET to disclose to the ABS any information held in the CCMS which is '*necessary for the creation and linkage of ECEC data sets by the ABS*', including information about children:
 - (i) who were included in the 2012 and 2015 AEDC; and
 - (ii) attended an approved (or formerly approved) child care service during the periods 2005 – 2015 (for children identified in the 2012 AEDC), and 2008 – 2015 (for children identified in the 2015 AEDC).

⁹ The SSRI Data provided to the ABS does not include CRNs.

- 
- (m) The CCMS PIC specifically identifies a child's age, gender, Indigenous status, and the number of hours of care provided to the child by the child care service. The ABS has also undertaken a review of the CCMS Data Items, and advises that all of the relevant information is *'necessary for the creation and linkage of ECEC data sets by the ABS'* for the purposes of the Project. On that basis, paragraph 168(1)(a) of the FAA and the CCMS PIC authorise:
 - (i) the disclosure of CCMS Data by AGDET to the ABS; and
 - (ii) by implication, the collection of that information by the ABS.

4.2 Collection of AEDC, SSRI, CCMS and NAPLAN Data

- (a) Under APP 3, the ABS must:
 - (i) only collect personal information (other than sensitive information) if it is reasonably necessary for, or directly related to, one or more of the ABS' functions or activities (APP 3.1);
 - (ii) only collect personal information by lawful and fair means (APP 3.5); and
 - (iii) only collect personal information directly from the relevant individual unless:
 - (A) the individual consents to the collection from someone else (APP 3.6(a)(i));
 - (B) the indirect collection is required or authorised by or under an Australian law (APP 3.6(a)(ii)); or
 - (C) it is unreasonable or impracticable to collect the information directly from the individual (APP 3.6(b)).
- (b) Under APP 3.3, the ABS can only collect sensitive information (such as health information, and Indigenous status) if:
 - (i) the individual consents, and the collection is reasonably necessary for, or directly related to, one or more of the ABS functions or activities; or
 - (ii) an exception in APP 3.4 applies, including (relevantly) where the collection is authorised or required by law.
- (c) The ABS advises that it has undertaken a review of the AEDC, SSRI, CCMS and NAPLAN Data Items, and considers that all of the information being collected is reasonably required by the ABS to undertake the data linkage and analysis processes and produce the desired data outputs. On that basis, we consider that for the purposes of APP 3.1, the collection of that information (insofar as it is personal information other than sensitive information) will be reasonably necessary for, or directly related to, the ABS functions under subsection 9(1) of the Census and Statistics Act.
- (d) We also consider that the collection of the AEDC, SSRI, CCMS and NAPLAN Data will be consistent with the ABS' obligations under APPs 3.3, 3.5 and 3.6, on the basis that:
 - (i) the collection of each of these datasets (including sensitive health and Indigenous status information contained in the datasets) by the ABS will be authorised under subsection 9(1) of the Census and Statistics Act;
 - (ii) the collection of the SSRI data items listed in Schedule 5 will be authorised under subsection 202(2C) of the SSA Act and the DSS PID;
 - (iii) the collection of the CCMS Data will be authorised under paragraph 168(1)(a) of the FAA and the CCMS PIC; and
 - (iv) it would be unreasonable or impracticable in the circumstances for the ABS to collect the AEDC, SSRI, CCMS and NAPLAN Data directly from the individuals concerned.

- (e) A 'collection' may also take place when an agency generates personal information from other data it holds. In the context of the Project, this will include data contained in the linkage input, record concordance and linked analysis datasets, which will be created by the ABS during the data linkage and data integration processes. We consider the collection of this information to be permitted under APP 3.1, on the basis that it is reasonably necessary for, or directly related to, the ABS' functions under subsection 9(1) of the Census and Statistics Act.

4.3 Meeting community expectations

- (a) As discussed above in section 2.4, the ability to rely on legislative authority for the collection and handling of personal information may ensure that a particular activity strictly complies with the APPs, but it does not necessarily mean that the activity will align with community expectations.
- (b) Further, a project's success can be impacted by the degree of community support and confidence in an agency's information handling practices. In this regard, transparency around the collection and handling of personal information can assist in building public trust. Conversely, insufficient transparency can undermine public confidence not only in relation to a particular project, but in relation to the agency as a whole.
- (c) The OAIC's *Community Attitudes to Privacy Survey 2017* found that nearly half (49%) of Australians are 'uncomfortable' with government agencies sharing their personal information with other agencies.¹⁰ Further, while 46% of survey respondents indicated they were comfortable with government agencies using their personal information for the purposes of 'research, improved services or policy development', a substantial proportion (approximately 30%) were uncomfortable with this idea.
- (d) Having regard to the number of datasets, and the volume and nature of data, there is a risk that members of the public – and particularly parents – may have concerns about the Project. For example:
 - (i) The VDC Data will comprise a rich dataset about individuals. Although the information will be de-identified, some members of the public may nevertheless hold concerns about the risk of re-identification, particularly in light of media attention relating to data breaches. For example, in September 2016 University of Melbourne academics notified the Department of Health that it was possible to figure out some service provider ID numbers from (what was supposed to have been de-identified) Medicare Benefits Schedule and Pharmaceutical Benefits Scheme datasets published on data.gov.au. These types of concerns may be heightened where the information is sensitive information, and information about children generally.
 - (ii) The 2016 Census raised concerns amongst privacy interest groups and the public generally in relation to the linking of Census data with other data collections. Although the 2011 Population Census Data does not include names or addresses, similar concerns could still be raised in relation to linking that data with other large datasets that do include names and other personal identifiers.
- (e) A loss of public confidence could adversely impact the ABS' reputation and the willingness of individuals to participate in other ABS projects. It could also have a broader impact on the willingness of individuals to participate in other government projects, due to concerns that their information could be linked with other information and used for purposes unrelated to the purposes for which it was provided.
- (f) To build public confidence and buy-in into the Project, it will be important to:
 - (i) be as transparent as possible in relation to the Project;
 - (ii) demonstrate how the public benefits of the Project will materially outweigh the potential privacy risks; and
 - (iii) explain how privacy risks will be mitigated.

¹⁰ *Community Attitudes to Privacy Survey 2017*, page 11.

4.4 Collection transparency

- (a) A key privacy right is for individuals to be made aware of what personal information is going to be collected about them, and for what purposes. In particular, APP 5 requires the ABS to take reasonable steps to notify or ensure an individual is made aware of certain matters at the time it collects their personal information or, if that is not practicable, then as soon as practicable after the collection. The matters in APP 5.2 relevantly include the following:
- (i) the fact that the ABS collects or has collected personal information from someone other than the individual (APP 5.2(b));
 - (ii) if the collection of the personal information is required or authorised by law – the fact that the collection is so required or authorised (including the name of the relevant law) (APP 5.2(c));
 - (iii) the purposes for which the ABS collects the personal information (APP 5.2(d));
 - (iv) information about how the individual may access and seek correction of their personal information, and how to make a complaint about a breach of the APPs (APPs 5.2(g) and (h)).
- (b) What constitutes 'reasonable steps' will depend on the circumstances. Relevant factors include:
- (i) the sensitivity of the information – more rigorous steps may be required when 'sensitive information', or information of a sensitive nature, is collected;
 - (ii) the possible adverse consequences for an individual as a result of collection; and
 - (iii) the practicability, including time and cost involved.¹¹
- (c) There is a risk that individuals may not be aware or would not reasonably expect the ABS to collect information about themselves or their children, particularly in the context of data linking. For example, parents of children may reasonably expect NAPLAN Data to be used for research relating to the national school curriculum, but not necessarily for broader research purposes (and noting also that parents in different jurisdictions may have been provided with different information regarding the use of NAPLAN Data). Centrelink customers may be even less likely to be aware or expect information provided to DHS for social security payment purposes may be used for research relating to childhood education.
- (d) Further, the legislative authority for collecting personal information may not be apparent or clear to individuals, particularly in relation to the scope of the ABS' powers to collect information under section 9 of the Census and Statistics Act.¹²
- (e) The ABS proposes to raise public awareness about the Project by including information on the Project on the ABS website, and on the National Statistical Service (**NSS**) register¹³ (or its future replacement).
- (f) Another way in which the ABS could promote transparency in relation to its data integration activities is through the ABS Privacy Policy. We understand that the Privacy Policy is currently under review, and is expected to be updated following that review. The inclusion of information in the Privacy Policy about the ABS data integration activities would also promote compliance with the ABS' obligations under APPs 1.3 and 1.4 (content of APP privacy policy).

¹¹ APP Guidelines, paragraph 5.4.

¹² For example, a number of submissions to the Senate Economics Reference Committee's *Inquiry into the 2016 Census* queried the interpretation and scope of the ABS' power under section 12 of the Census and Statistics Act to collect names and addresses. See the discussion at paragraphs 4.44 – 4.48 of the Senate Committee's report, *2016 Census: Issues of trust (November 2016)*, available at http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Economics/2016Census/Report.

¹³ <http://www.nss.gov.au/nss/home.NSF/pages/Data%20Integration:%20Project%20Register%20-%20Help>.

- 
- (g) There may also be opportunities to raise awareness about the handling of personal information in relation to the Project via Project stakeholders (for example, privacy statements and other public facing information communicated by or through those stakeholders).

Recommendation 1: To promote transparency and public debate about the benefits of, or any privacy concerns relating to, the Project, the ABS should ensure that:

- its communication strategy (which may be delivered through a range of activities and media channels) reaches as much of the Australian community as possible, and particularly parents with children; and
- easily accessible information is made available which covers the following matters:
 - why the Project is important;
 - what information will be collected and used;
 - the scope of the statistical and research purposes for which the integrated VDC data may be used;
 - an explanation of the legislative authority relied upon by the ABS to collect and use data (including the use of 2011 Population Census Data) for the purposes of the Project;
 - how the data linkage will occur; and
 - how information will be protected.

The ABS may wish to consult and/or coordinate with other Project stakeholders (e.g. education departments) in relation to the development and implementation of appropriate communication measures.

Recommendation 2: Further to Recommendation 1:

- the project description on the National Statistical Service (**NSS**) Data Integration Project register (or its replacement) should include information about the nature of the information that will be collected and used by the ABS; and
- the ABS website should include prominent and clear information about the Project, including a link to the NSS Data Integration Project register (or its replacement).

Recommendation 3: As part of the review of the ABS Privacy Policy, consideration should be given to including information about the collection of data from Commonwealth, State and Territory government bodies, and the use of that information in data integration projects.

4.5 Quality of personal information

- (a) APP 10 requires the ABS to take reasonable steps to ensure that:
- (i) the personal information collected by the ABS is accurate, up to date and complete (APP 10.1); and
 - (ii) the personal information that the ABS uses or discloses is (having regard to the purpose of the use or disclosure) accurate, up to date, complete and relevant (APP 10.2).
- (b) What constitutes 'reasonable steps' depends on the circumstances. Relevant factors may include the sensitivity of the information, the nature of the entity holding the information,



and the possible adverse consequences for an individual if the quality of personal information is not ensured.¹⁴

- (c) The ABS will rely on DSS, AGDET and NAPLAN Data Custodians as 'trusted sources' of AEDC, SSRI, CCMS and NAPLAN data. These data custodians provide the ABS with information about the quality of the data, and the ABS obtains further information from those data custodians if required after it has inspected the data.
- (d) Data linkage processes during the course of the Project will also be subject to quality assurance.
- (e) In the circumstances, we do not consider the Project raises any material APP 10 compliance risks.

¹⁴ APP Guidelines, paragraph 10.6.

5. Data Integration

- (a) Under APP 6.1, personal information that has been collected by the ABS for a particular purpose (the 'primary purpose') must not be used or disclosed for another purpose (the 'secondary purpose') unless the individual has given their consent, or a permitted circumstance specified in APP 6.2 applies. These exceptions include where:
 - (i) the individual would reasonably expect the agency to use or disclose the information for the secondary purpose, and the secondary purpose is:
 - (A) if the information is 'sensitive information' – directly related to the primary purpose; or
 - (B) if the information is not sensitive information – related to the primary purpose (APP 6.2(a)); or
 - (ii) the use or disclosure is required or authorised by or under an Australian law (APP 6.2(b)).
- (b) The data linkage and integration processes will involve a use of personal information (including sensitive information) included in AEDC, NECECC, SSRI, CCMS, NAPLAN and 2011 Population Census Data, as well as information contained in the linkage input, record concordance and linkage analysis datasets.

Use of AEDC, SSRI, CCMS and NAPLAN Data

- (c) We consider that the use of AEDC, SSRI, CCMS and NAPLAN Data will be in accordance with APP 6, on the basis that:
 - (i) the AEDC, SSRI, CCMS and NAPLAN Data will only be used for the primary purpose for which it is collected by the ABS (i.e. to prepare the integrated VDC dataset for this Project), and at this time is not proposed to be used for any other purpose;¹⁵
 - (ii) the use of SSRI Data for the purposes of the Project is authorised under subsection 202(2C) of the SSA Act and the DSS PID; and
 - (iii) the use of CCMS Data for the purposes of the Project is authorised under paragraph 162(2)(e) of the FAA Act.¹⁶

Use of NECECC and 2011 Population Census Data

- (d) The 2011 NECECC Data was collected by the ABS under subsection 9(1) of the Census and Statistics Act. The 2011 Population Census Data was collected by the ABS under section 8 of the Census and Statistics Act. .
- (e) Subsection 12(1) of the Census and Statistics Act provides that the '*Statistician shall compile and analyse the statistical information collected under this Act and shall publish and disseminate the results of any such compilation and analysis, or abstracts of those results*'.
- (f) The ABS advises that:
 - (i) Population Census and NECECC Data is collected for statistical purposes under the Census and Statistics Act, and is not limited to the use of data for a specific statistical compilation;
 - (ii) data integration is one method of compiling statistics by combining data from different sources, through linking information common to both datasets; and
 - (iii) therefore, in accordance with subsection 12(1) of the Census and Statistics Act, the NECECC and 2011 Population Census Data will be used to compile and

¹⁵ If the relevant data custodians agree in the future to the use of datasets for purposes other than this Project, that proposed use will be subject to a separate privacy assessment.

¹⁶ Paragraph 162(2)(e) of the FAA Act authorises a person to use protected information for the purpose for which the information was disclosed to the person under sections 167 or 168, including the purpose set out in a PIC made under paragraph 168(1)(a) of the FAA Act.

analyse a new statistical dataset (the integrated VDC Data), the results (or abstracts of results) of which will be published in a de-identified form.

- (g) On that basis:
- (i) the 2011 Population Census and NECECC Data will be used for the primary purpose for which it was collected by the ABS, and therefore will be in accordance with APP 6.1; and
 - (ii) even if this were to be characterised as a use for a secondary purpose:
 - (A) subsection 12(1) of the Census and Statistics Act would provide a legislative authority for the ABS to use statistical information collected under that Act to compile statistics on any matter covered by the Census and Statistics Regulation (including education and social welfare services¹⁷); and
 - (B) the use of NECECC and 2011 Population Census Data for the purposes of data linkage and integration in relation to the Project, will therefore be 'authorised by law' and permitted under APP 6.2(b).
- (h) In the interests of transparency, we recommend that the ABS ensure there is clear and publicly available information which explains the legislative basis on which data (including in particular 2011 Population Census data) will be used in relation to the Project (see **Recommendation 1**).

Risk of use of data for other secondary purposes

- (i) The extent to which the relevant datasets will be used (i.e. accessed and handled by staff) within the ABS will be limited by 'functional separation' procedures which ensure identifying information is used only for the purposes of data linkage, and that staff only access information on a 'need to know' basis according to the following key data linkage roles:
- (i) **Librarian:** The Librarian performs processes such as assigning a random identifier for each record, anonymisation (i.e. de-identification, such as scrambling individual's names) and other data preparation activities. The Librarian also creates concordances between original record IDs for an individual, and newly created randomised IDs, which are provided to the Assembler to enable the linking of different datasets.
 - (ii) **Linker:** The Linker performs the linkage of the datasets being brought together. Access is limited to the fields required for linking.
 - (iii) **Assembler:** The Assembler creates the linked analysis file. The Assembler has access to the record concordances produced as part of the linkage process, and analysis variables, allowing them to produce the linked analysis file. The Assembler does not have access to identifying variables used for linking.
 - (iv) **Analyst:** The Analyst conducts analysis on the linked analysis file, and produces statistical data for release. Access is limited to the data needed for analytical purposes, and does not include identifying information or information that might allow content data to be linked back to identifying information.
- (j) We consider the functional separation approach to be a positive measure to mitigate risks relating to the use of the datasets (including the final VDC dataset) for secondary purposes, including future 'function creep' risks. We also note that any proposed use of the final VDC dataset for another (or additional) project will, in accordance with the ABS' governance processes, be subject to a separate project approval process that includes an assessment of privacy risks.

¹⁷ Census and Statistics Regulation, items 14 and 45 in regulation 13.

6. Availability of VDC data

- (a) Sections 12 and 13 of the Census and Statistics Act require the ABS to ensure that information is not released in a way that is likely to enable an individual to be identified.
- (b) A de-identified microdata file of the VDC Data will be made available to authorised users (which may include persons external to the ABS) via the ABS DataLab. The DataLab is a data analysis solution that allows users to undertake interactive analysis of microdata.
- (c) Data that has been successfully de-identified is not 'personal information' for the purposes of the Privacy Act. Personal information is de-identified if the information is no longer about an identifiable individual or an individual who is reasonably identifiable.¹⁸
- (d) The volume and richness of the data, together with the large number of variables (including detailed geographical information) in the VDC dataset raises a potentially significant statistical disclosure risk. This requires the VDC dataset to be adequately de-identified, and adequate measures to be implemented to mitigate the risk of re-identification (including as a result of the data user combining the VDC dataset with other datasets available to the user).
- (e) The following measures will be applied to reduce the risk of releasing data that has not been properly de-identified, and the risk of re-identification:¹⁹
 - (i) The VDC dataset will remain in the secure ABS security environment.
 - (ii) Before the VDC dataset is made available, the data is subject to a number of 'confidentialisation' methods (e.g. collapsing response categories, top and bottom coding²⁰), and review to ensure that potentially identifiable information is not disclosed.
 - (iii) Access to the DataLab is restricted to authorised users. To become an authorised user, the researcher must:
 - (A) sign legally binding undertakings that they will only use the data for the approved statistical research purpose, will not attempt to re-identify or match the data with any other information, and will not disclose the data to any other person without the approval of the Australian Statistician;²¹
 - (B) have sufficient statistical knowledge and experience; and
 - (C) only access the data for the approved statistical purposes.
 - (iv) Microdata is released at the discretion of the ABS following an assessment process, which includes assessing the risk of a particular user being able to re-identify the data. Further, all analytical outputs are checked by the ABS before being provided to the authorised researcher.
 - (v) The DataLab can only be accessed at a secure on-site ABS location, or as part of the ABS virtual DataLab. System controls prevent a researcher from being able to download or print microdata, or merging other datasets with the microdata. Access may also be limited to a subset of the overall available data, based on the ABS assessment of the researcher's access request (including what data they wish to access and why).
 - (vi) Access to the DataLab is regularly audited and monitored.

¹⁸ Privacy Act, s 6(1): definition of 'de-identified'

¹⁹ [http://www.abs.gov.au/websitedbs/D3310114.nsf/home/CURF:+About+the+ABS+Data+Laboratory+\(ABSDL\)](http://www.abs.gov.au/websitedbs/D3310114.nsf/home/CURF:+About+the+ABS+Data+Laboratory+(ABSDL)).

²⁰ Top and bottom coding is a method used to ensure microdata is adequately de-identified, and involves replacing certain numerical variables to mask data that is potentially identifying.

²¹ It is an offence under s 19(3) of the Census and Statistics Act for a person to fail to comply with an undertaking given under s 13(2)(c).

7. Other Privacy Considerations

7.1 Anonymity and pseudonymity

- (b) APP 2.1 requires the ABS to give individuals the option of not identifying themselves, or of using a pseudonym, when dealing with an APP entity in relation to a particular matter, unless:
 - (i) the ABS is required or authorised by or under an Australian law, or by a court/tribunal order, to deal with individuals who have identified themselves; or
 - (ii) it is impracticable for the ABS to deal with individuals who have not identified themselves, or who have used a pseudonym.
- (c) APP 2 has limited practical application in relation to the Project, as the ABS will not be dealing directly with individuals to whom the relevant datasets relate. The ABS also requires information about identifiable individuals in order to perform the data linkage between the relevant datasets (and noting that the final VDC dataset will be de-identified).
- (d) We therefore do not consider the Project raises any material APP 2 compliance risks.

7.2 Data security

- (a) APP 11 requires the ABS to take 'such steps as are reasonable in the circumstances' to maintain the security of the personal information it holds. This includes risks relating to unauthorised disclosure by ABS staff, or unauthorised access by external parties for malicious purposes such as theft or identity fraud.
- (b) The types of security measures that an agency could implement to meet its APP 11 obligations may include the following:
 - (i) governance, culture and training;
 - (ii) internal practices, procedures and systems;
 - (iii) ICT security;
 - (iv) access security;
 - (v) contractual obligations and arrangements with third party providers;
 - (vi) data breach response plan;
 - (vii) physical security;
 - (viii) destruction and de-identification of data; and
 - (ix) compliance with data security standards.²²
- (c) The ABS advises that the following security strategies will apply in relation to the Project, and are measures which have been informed by a privacy and security risk assessment process undertaken by the ABS:
 - (i) the ABS' IT systems comply with IT security arrangements set out in the Australian Signals Directorate's (**ASD**) *Australian Government Information Security Manual*;
 - (ii) access to all ABS premises is strictly controlled in accordance with the *Protective Security Policy Framework (PSPF)*;
 - (iii) upon appointment, all ABS staff are:
 - (A) subject to security checks, including security clearances when required; and
 - (B) are required to sign an *Undertaking of Fidelity and Secrecy* under section 7 of the Census and Statistics Act – it is an offence under section 19 of that Act for an ABS officer to disclose information obtained under the Act;

²² APP Guidelines, paragraph 11.8. See also the OAIC's *Guide to securing personal information*.

- 
- (iv) all ABS staff are required to undertake compulsory biennial refresher training in relation to their security and confidentiality responsibilities;
 - (v) the ABS internet gateway is annually reviewed for compliance against ISM requirements by an independent ASD accredited Information Security Registered Assessors Program (**IRAP**) assessor;
 - (vi) regular PSPF risk reviews are undertaken in relation to ABS security arrangements to ensure they remain effective;
 - (vii) there is an ongoing program of security audits and reviews of computer systems and the physical environment;
 - (viii) access to data is limited on a need to know basis according to role-based access controls, and only internal ABS staff in the DLC will have access to unit record data;
 - (ix) ABS staff with access to name, address and other linking variables will not have access to any analytical variables; and
 - (x) the ABS has a data breach response plan in place (which includes notification to affected individuals and the OAIC), and a Privacy Management Plan is under development.
- (d) We consider the above measures will constitute 'reasonable steps' for the purposes of APP 11.

7.3 Data retention

- (a) The unnecessary retention of personal information that is no longer required for a function or activity can raise potential risks in relation to the use of that information for secondary purposes, as well as data security risks.
- (b) At the end of the Project (anticipated in the last quarter of 2017), a decision will be made by the ABS, DSS, AGDET and the NAPLAN Data Custodians as to whether to retain the linked VDC Data. The source datasets will be deleted/destroyed when no longer needed for the purposes of creating the linked VDC Data, except where approval has been received by the data custodians to retain the data for other purposes (as is the case for the AEDC data).
- (c) Any proposal to use the VDC Data for another purpose (including the further linking of the VDC Data with additional datasets) will be subject to a separate approval process, including a PIA.

7.4 Data access and correction

- (a) Under APP 12, individuals must be provided with access to their own personal information, unless an exception applies (for example, if the agency is required or authorised to refuse access under another Commonwealth law that provides for access to documents, such as the *Freedom of Information Act 1982* (Cth)).
- (b) APP 13 requires agencies to take reasonable steps, on request by an individual, to correct any personal information they hold to ensure it is accurate, relevant, up to date, complete and not misleading, having regard for the purpose for which it is held.
- (c) The ABS has existing processes in place to enable individuals to access and seek correction of their own personal information. We consider this will satisfy the ABS' obligations under APPs 12 and 13, noting that:
 - (i) under APP 12, an agency can refuse to give an individual access to their own personal information where the agency is required or authorised to do so under the *Freedom of Information Act 1982* (Cth) (**FOI Act**);

- 
- (ii) pursuant to subsection 7(2) and Part II, Division 2 in Schedule 2 of the FOI Act, the ABS is exempt from the operation of the FOI Act in relation to documents containing information collected under the Census and Statistics Act, which means that under APP 12.2 the ABS is not required to give an individual access to such information.

7.5 Privacy Policy

- (a) APP 1.3 requires the ABS to have a clearly expressed and up to date privacy policy about its management of personal information, which addresses the matters specified in APP 1.4. These matters relevantly include:
 - (i) the kinds of personal information that the ABS collects and holds (APP 1.4(a));
 - (ii) how the ABS collects and holds personal information (APP 1.4(b)); and
 - (iii) the purposes for which the ABS collects, holds, uses and discloses personal information (APP 1.4(c)).
- (b) We refer to Recommendation 3 in relation to addressing data integration activities in the ABS Privacy Policy.

Schedule 1 – Detailed information flows

Step	Description	PIA Information Flow
A. Data Sourcing		
AEDC Data (2012)		
A1.	<p>AGDET provides the ABS with two separate files: a password protected encrypted linkage dataset, and a de-identified analysis dataset relating to the AEDC Data.</p> <ul style="list-style-type: none"> The de-identified analysis dataset contains personal information about children provided by their teachers. It also includes sensitive information, such as their Indigenous status and certain health information (e.g. whether they child has been assessed as having 'special needs' because they are blind, have asthma, are autistic, etc). The password protected encrypted linkage dataset is downloaded from a Secure Exchange website by the ABS' Administrative Data Acquisition Unit (ADAU), and then securely transferred to a limited access 'linkage dataset' folder controlled by the DLC. The de-identified analysis dataset is downloaded from a Secure Exchange website by the ADAU, and then securely transferred to a limited access 'analysis dataset' folder controlled by the DLC. 	<p>ABS collects personal information (including sensitive information) from AGDET about children, for the purposes of VDC data integration.</p> <ul style="list-style-type: none"> ➤ Statistical information is collected under s 9(1) of the Census and Statistics Act.
NECECC Data (2011)		
A2.	<p>NECECC Data is held by the Education and Training Statistics Branch in the ABS. The NECECC Data is provided to the ABS DLC as two separate linkage and analysis files.</p>	<p>By providing the NECECC Data to another area within the ABS, the ABS uses personal information (including sensitive information) about children.</p>
2011 Population Census Data		
A3.	<p>DLC currently holds a de-identified version of 2011 Population Census Data (which does not contain any names or addresses) in DLC's secure holdings.</p>	<p>N/A – information in the de-identified 2011 Population Census dataset (including information about parents and children) is already held by the DLC. There is no new transfer or provision of this information from the Census Dissemination Branch (which holds 2011 Population Census Data) to DLC for the purposes of the Project.</p>

Step	Description	PIA Information Flow
SSRI Data (2005 – 2015)		
A4	<p>AGDET provides the ABS with the SSRI Data on two separate linkage and analysis datasets.</p> <ul style="list-style-type: none"> • The SSRI Data is initially provided by DSS to AGDET. • AGDET securely uploads the linkage and analysis datasets to the ABS Secure Deposit Box. <ul style="list-style-type: none"> ○ The de-identified SSRI dataset includes information originally collected by DHS from Centrelink customers for the purposes of administering payments under the social security law and family assistance law. ○ The de-identified analysis dataset contains information about children, their parents/carers and the parent/carer's spouse. • The datasets are then securely transferred by the ADAU to limited access 'linkage dataset' and 'analysis dataset' folders controlled by the DLC. 	<p>ABS collects personal information from AGDET about children, their parents/carer and the parent/carer's spouse, for the purposes of VDC data integration.</p> <ul style="list-style-type: none"> ➤ Statistical information is collected under s 9(1) of the Census and Statistics Act. ➤ The collection of SSRI Data is also authorised by the DSS PID made under s 202(2C) of the SSA Act.
CCMS Data (2005 – 2015)		
A5	<p>AGDET provides the ABS with the CCMS Data on two separate linkage and analysis datasets.</p> <ul style="list-style-type: none"> • The CCMS Data is initially provided by DSS to AGDET. • AGDET securely uploads the linkage and analysis datasets to the ABS Secure Deposit Box. <ul style="list-style-type: none"> ○ The de-identified analysis CCMS Data includes information originally collected from child care service providers in relation to child enrolment and attendance for FTB, child care benefit and child care rebate purposes. ○ The analysis datasets contains de-identified personal information (including sensitive information) about children. • The datasets are then securely transferred by the ADAU to limited access 'linkage dataset' and 'analysis dataset' folders controlled by the DLC. 	<p>ABS collects personal information (including sensitive information) from AGDET about children, for the purposes of VDC data integration.</p> <ul style="list-style-type: none"> ➤ Statistical information is collected under s 9(1) of the Census and Statistics Act. ➤ The collection of CCMS Data is also authorised by the CCMS PIC, made under s 168(1)(a) of the FAA Act.
NAPLAN Data (Year 3 results for 2014, 2015 & 2016)		
A6	<p>Each NAPLAN Data Custodian provides the ABS with NAPLAN Data.</p> <ul style="list-style-type: none"> • The datasets are uploaded by each NAPLAN Data Custodian to the ABS Secure Deposit Box, and then securely transferred to a limited access 'linkage dataset' and 'analysis dataset' files 	<p>ABS collects personal information (including sensitive information) about children, for the purposes of VDC data integration.</p> <ul style="list-style-type: none"> ➤ Statistical information is collected under s

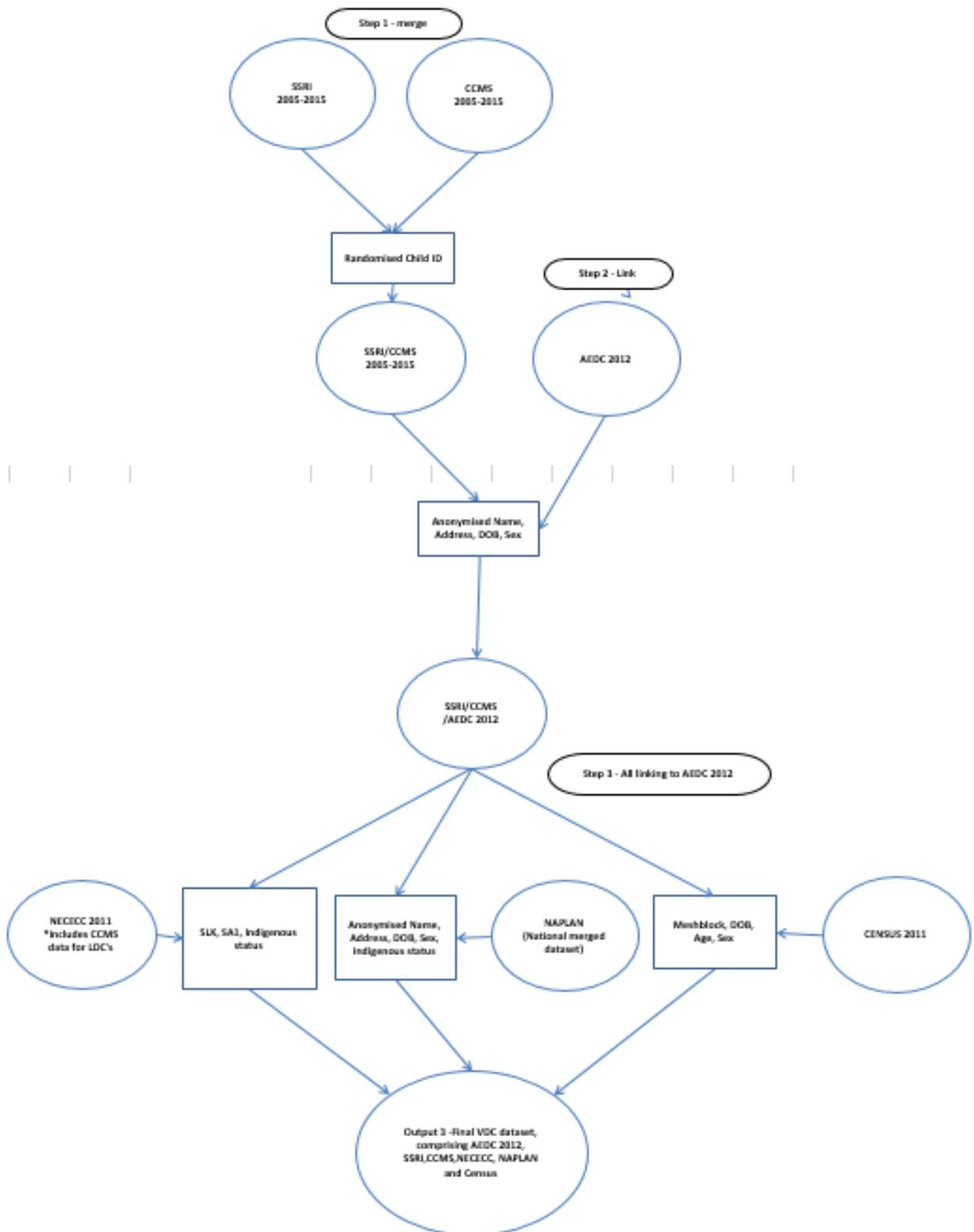
Step	Description	PIA Information Flow
	<p>controlled by the DLC.</p> <ul style="list-style-type: none"> Each jurisdiction's NAPLAN Data is provided as two separate datasets: a linkage and an analysis dataset. The de-identified NAPLAN Data includes personal information and sensitive information about children (Year 3 students). 	9(1) of the Census and Statistics Act.
B. Data Linkage		
B1.	Stage 1 Linkage	
B1.1	<p>The ABS Librarian:</p> <ul style="list-style-type: none"> receives the SSRI and CCMS linkage datasets, and the AEDC linkage datasets; combines (merges) the CCMS and SSRI linkage datasets using the randomised unique child identifier (assigned by AGDET); creates linkage input datasets in relation to the SSRI/CCMS and AEDC datasets, with newly assigned randomised record IDs (both in relation to children, as well as the child's parents/carers and their spouses); applies an anonymisation process to randomly scramble individuals' names in the relevant data files; addresses are geocoded to Mesh Block and higher geographies, and removed from the linkage input file; creates a record concordance, which contains the correspondence between the original record IDs (in relation to each individual) and the new linkage IDs; creates analytical variables using the SSRI/CCMS and AEDC Data; provides the linkage input datasets to the ABS Linker (one each for the SSRI/CCMS Data and the AEDC 2012 Data); and securely transfers the record concordance and analytical variables file to the ABS Assembler. 	<p>ABS:</p> <ul style="list-style-type: none"> uses personal information about children contained in the SSRI, CCMS and AEDC datasets, and personal information about the child's parents/carers and the parent/carer's spouse contained in the SSRI Data, to create the linkage input datasets, the record concordance and analytical variables; and collects personal information about children, parents/carers and their spouses in the form of the linkage input datasets and record concordances for the purposes of VDC data integration.

Step	Description	PIA Information Flow
B1.2	<p>The ABS Linker:</p> <ul style="list-style-type: none"> links the SSRI/CCMS and AEDC Data (input datasets), using the child's anonymised name, geocode, DOB and sex as the key linking variables; as a result of the linkage process – produces concordances (which contain the linkage IDs and quality indicators, and illustrates a direct link between the random IDs in each dataset) for the SSRI/CCMS/AEDC 2012 Data; and securely transfers the record concordance to the ABS Assembler. 	<p>ABS:</p> <ul style="list-style-type: none"> uses personal information contained in the input datasets (SSRI/CCMS and AEDC Data) to create a linkage output file; and collects personal information in the form of linkage IDs included in the record concordance, <p>for the purposes of VDC data integration.</p>
B1.3	<p>The ABS Assembler creates linked analysis datasets using:</p> <ul style="list-style-type: none"> the analytical variables and record concordance provided by the Librarian; and the record concordance provided by the Linker, <p>which (collectively) contain the original IDs and the randomised IDs, and produces an output linked dataset (i.e. the linked SSRI/CCMS/AEDC 2012 Data file).</p> <p>Another randomised ID is created to replace the initial randomised ID during this process.</p>	<p>ABS:</p> <ul style="list-style-type: none"> uses personal information to create linked analysis datasets and an output linked dataset; and collects personal information in the form of the linked analysis datasets and the output linked dataset, <p>for the purposes of VDC data integration.</p>
B2.	Stage 2 Linkage	
B2.1	<p>The ABS Librarian:</p> <ul style="list-style-type: none"> receives the linked SSRI/CCMS/AEDC 2012 Data file (output from Stage 1); combines the NAPLAN Data provided by each NAPLAN Data Custodian into a single NAPLAN dataset using the child's name or SLK, address or Mesh Block, DOB, sex and school ID; receives the 2011 Population Census Data and NECECC Data; applies an anonymisation process to name information in the relevant datasets; creates linkage input datasets in relation to the SSRI/CCMS/AEDC 2012 Data, NAPLAN Data, NECECC Data and 2011 Population Census Data, with newly assigned record linkage IDs; addresses are geocoded and removed from the linkage input dataset; creates a record concordance, which contains the correspondence between the original record 	<p>ABS:</p> <ul style="list-style-type: none"> uses personal information contained on the SSRI/CCMS/AEDC 2012, NAPLAN and 2011 Population Census Data files; and collects personal information in the form of linkage input datasets and record concordances, <p>for the purposes of VDC data integration.</p>

Step	Description	PIA Information Flow
	<p>IDs (in relation to each individual) and the new linkage IDs;</p> <ul style="list-style-type: none"> creates analytical variables for the Assembler; provides the linkage input datasets to the ABS Linker (one each for the SSRI/CCMS/AEDC 2012 Data, NAPLAN Data, NECECC Data and 2011 Population Census Data); and securely transfers the record concordance and analytical variables file to the Assembler. 	
B2.2	<p>The ABS Linker:</p> <ul style="list-style-type: none"> links the SSRI/CCMS/AEDC 2012 Data to the NECECC, NAPLAN and 2011 Population Census Data, using certain key linkage variables such as the child's anonymised name, geocode, DOB and sex (however in the case of 2011 Population Census data, the 'Bronze Standard' does not use name and address information in the linkage process); as a result of the linkage process – produces concordances (which contain the linkage IDs and quality indicators, and illustrates the direct links between the random IDs in each dataset) for the combined VDC data set comprising SSRI, CCMS, AEDC 2012, NECECC, NAPLAN and 2011 Population Census Data; and provides the concordance (i.e. output linkage dataset) to the ABS Assembler. 	<p>ABS:</p> <ul style="list-style-type: none"> uses personal information contained in the input datasets (SSRI/CCMS/AEDC 2012 Data to the NECECC, NAPLAN and 2011 Population Census Data) to create a output linkage dataset; and collects personal information in the form of linkage IDs contained in the linkage output dataset, <p>for the purposes of VDC data integration.</p>
B2.3	<p>The ABS Assembler creates a final linked analysis dataset for the combined VDC dataset using:</p> <ul style="list-style-type: none"> the analytical variables and record concordance received from the Librarian; and the record concordance received from the Linker. <p>Another randomised ID is created to replace the previous randomised ID.</p>	<p>ABS:</p> <ul style="list-style-type: none"> uses personal information to create a linked analysis dataset; and collects personal information in the form of the linked analysis dataset, <p>for the purposes of VDC data integration.</p>
C. Data availability		
C1.	A de-identified microdata file of the VDC Data will be accessible to authorised users via the ABS DataLab.	No personal information will be disclosed.

Schedule 2 – Data Linkage Process Diagram

The following diagram has been prepared by the ABS to illustrate the linkage process in relation to the various datasets for this Project.



Schedule 3– Data Linkage Process: Linking and Analysis Variables

Dataset	Linking Variables	Analysis Variables
CCMS 2005 - 2015	<ul style="list-style-type: none"> ▪ Unique child identification code ▪ Unique customer identification code ▪ Anonymised child name ▪ Geocodes ▪ Child and Parent/Carer DOB ▪ Sex 	<ul style="list-style-type: none"> ▪ Child/family demographics ▪ Child care provider characteristics ▪ Child care enrolment and attendance indicators
SSRI 2005 - 2015	<ul style="list-style-type: none"> ▪ Unique child identification code ▪ Unique customer identification code ▪ Anonymised child name ▪ Geocodes ▪ Child and Parent/Carer DOB ▪ Sex 	<ul style="list-style-type: none"> ▪ Child and family characteristics, including 1 or 2 parent household, number of children in the household, age of household members ▪ Transfer payment types ▪ Employment status of parent(s) ▪ Duration of payments ▪ Healthcare card status
AEDC 2012	<ul style="list-style-type: none"> ▪ Anonymised school name or School ID ▪ Anonymised child name ▪ Geocodes (for child and school location) ▪ Child DOB ▪ Child Indigenous status ▪ SLK (derived by ABS from name, DOB and sex) ▪ SA1 (ABS geocoded) ▪ Mesh Block (ABS geocoded) 	<ul style="list-style-type: none"> ▪ Child demographics ▪ Information about teacher and school ▪ AEDC instrument variables
NECECC 2011	<ul style="list-style-type: none"> ▪ SLK (derived from child's name, DOB and sex) ▪ Geocodes ▪ Indigenous status 	<ul style="list-style-type: none"> ▪ Child/family demographics ▪ Provider characteristics ▪ Enrolment and attendance indicators

Dataset	Linking Variables	Analysis Variables
2011 Population Census	<ul style="list-style-type: none"> ▪ Geocodes ▪ Child and Parent/Carer DOB ▪ Age ▪ Sex 	<ul style="list-style-type: none"> ▪ All standard 2011 Population Census output items, which geographic and socio-economic characteristics of people, families and dwellings²³
NAPLAN 2014 – 2016 ²⁴	<ul style="list-style-type: none"> ▪ Geocodes (for child and school location) ▪ Anonymised child's name (not available in all jurisdictions) ▪ SLK (where name cannot be provided) ▪ Child's DOB ▪ Sex ▪ Indigenous status ▪ Anonymised school name or School ID 	<ul style="list-style-type: none"> ▪ Child achievement and participation in literacy and numeracy at Year 3 (tests results in reading, writing, spelling, grammar and punctuation, and numeracy) ▪ Child and parents/carers characteristics ▪ Sector (government / non-government)

²³ See 2901.0 – *Census Dictionary, 2011*.

²⁴ All states and territories were invited to supply NAPLAN data. At the time of the report South Australia was yet to participate.

Schedule 4– NAPLAN Data Custodians

The following state and territory education departments and agencies and NAPLAN test administration authorities were invited to provide NAPLAN Data to the ABS for the purposes of the Project.

- NSW Department of Education / NSW Education Standards Authority
- Victorian Department of Education and Training / Victorian Curriculum and Assessment Authority
- Queensland Department of Education / Queensland Curriculum and Assessment Authority
- South Australian Department for Education and Child Development
- Western Australian Department of Education
- Tasmanian Department of Education
- Northern Territory Department of Education
- ACT Education Directorate

Schedule 5 – SSRI Data covered by DSS PID

The following is an extract of the DSS PID, which sets out the types of information (including personal information) covered by that instrument.

- CHILD

Centrelink customer CRN details of every person registered with a date of birth between 1.1.2005 and 31.12.2012 (including children who are deceased).

- a) Name (first, second and last names, alias / previously known as, name type, name change date, start and end date)
- b) Sex
- c) Date of birth (dd/mm/yyyy)
- d) Date of death (dd/mm/yyyy)
- e) Child disability status
- f) Country of birth
- g) Child carer relationship including in care episodes (start and end dates), primary care flag, foster care flag and shared care percentage

- CARER Demographics and Timestamps

- (a) Name (first, second, last name, also known as, maiden names, previously known as), Carer Name Type and Carer Name Change Date
- (b) Carer Address History and timestamps, address street number, street name, suburb, state, postcode, SA1 or historical equivalent, Community Identifier (4 letter identifier for community or outstation)
- (c) Address type (e.g. home term, temporary, postal, business).
- (d) Address start and end dates (dd/mm/yyyy)
- (e) Carer DOB (dd/mm/yyyy)
- (f) Carer Date of Death (dd/mm/yyyy)
- (g) Carer Sex
- (h) Carer Educational Level
- (i) Carer Preferred Language for correspondence
- (j) Interpreter language
- (k) Carer Indigenous status
- (l) Carer Country of Birth
- (m) Carer Benefit history: start date, end date, full/part/zero rate indicator, benefit type, end reason

- (n) Carer rent assistance/home ownership indicator
- (o) Carer Employment indicator, start and end dates
- (p) Carer annual income reported in order to determine eligibility for FTB
- (q) Carer has been eligible for low income card this calendar year
- (r) Carer has been identified as being vulnerable this calendar year, including vulnerability type and duration
- (s) Carer has been homeless this calendar year, including duration of homelessness

- SPOUSE of Carer Demographics and Timestamps
 - (a) Name (first, second, last name, also known as, maiden names, previously known as), Carer Name Type and Carer Name Change Date
 - (b) Spouse Address History and timestamps, address street number, street name, suburb, state, postcode, SA1 or historical equivalent, Community Identifier (4 letter identifier for community or outstation)
 - (c) Address type (e.g. home term, temporary, postal, business)
 - (d) Address start and end dates (dd/mm/yyyy)
 - (e) Spouse DOB (dd/mm/yyyy)
 - (f) Spouse Date of Death (dd/mm/yyyy)
 - (g) Spouse Sex
 - (h) Spouse Education Level
 - (i) Spouse Preferred Language for correspondence
 - (j) Interpreter language
 - (k) Spouse Indigenous status
 - (l) Spouse Country of Birth
 - (m) Spouse Benefit history: start date, end date, full/part/zero rate indicator, benefit type, and reason
 - (n) Spouse rent assistance/home ownership indicator
 - (o) Spouse Employment Indicator, start and end dates
 - (p) Spouse annual income reported in order to determine eligibility for FTB
 - (q) Spouse has been eligible for low income card this calendar year
 - (r) Spouse has been identified as being vulnerable this calendar year, including vulnerability type and duration
 - (s) Spouse has been homeless this calendar year, including duration of homelessness
- CARER SPOUSE relationship details
 - (a) Carer – Spouse Relationship Indicator specifically:
 - Relationship Status / type
 - Relationship Start Date
 - Relationship End date
 - (b) Carer Child primary carer count as at
 - 1 May 2012
 - 1 May 2015
 - (c) Carer child count as at
 - 1 May 2012
 - 1 May 2015
 - (d) Carer age youngest child in primary care as at
 - 1 May 2012
 - 1 May 2015

Schedule 6 – CCMS Data covered by CCMS PIC

The following is an extract of the DSS PID, which sets out the types of information (including personal information) covered by that instrument.

4 Information able to be disclosed

Any departmental records or information held by the Department in the Child Care Management System (CCMS) necessary for the creation and linkage of ECEC data sets by the ABS to allow for the analysis of linked ECEC data sets by the AIHW.

This information includes data relating to all children included in the 2012 and 2015 Australian Early Development Census (AEDC) in an approved, or formerly approved child care service during the period 2005 to 2015 (for children identified in the 2012 AEDC) and 2008 to 2015 (for children identified in the 2015 AEDC), which involves:

- (a) the service's operator name, entity type and CCMS unique identifier;
- (b) the service's address with geo-location information;
- (c) the number of weeks each calendar year the service operates;
- (d) the number of children provided care by the service;
- (e) the following characteristics of the children provided care by the service;
 - i. child's age, gender, indigenous status;
 - ii. statistical linkage key;
 - iii. number of hours the child was provided care by the service; and
 - iv. fees charged by the service for the child, including the total amount and the amount reflecting any reductions of child care benefit, child care rebate and other childcare subsidies; and
- (f) any other matter necessary for the purpose outlined in section 3 above.

Schedule 7 – Summary of the APPs

1. Collection of personal information

1.1 Anonymity and pseudonymity

Wherever it is lawful and practicable in the circumstances, APP 2 requires Commonwealth agencies to give individuals the option of interacting anonymously or by using a pseudonym.

1.2 Collection necessity

APP 3.1 requires Commonwealth agencies to collect personal information only where it is necessary for, or directly related to, a purpose that is directly related to a function or activity of the agency.

1.3 Collecting sensitive personal information

APPs 3.3 and 3.4 limit the circumstances in which Commonwealth agencies can collect 'sensitive information'. Some of the permitted circumstances include where the individual has given consent, where collection is required or authorised by or under an Australian law, or (where it is unreasonable or impracticable to obtain the individual's consent) the collection is necessary to lessen or prevent a serious threat to the life, health or safety of an individual.

1.4 Collection methods – lawful, fair and not intrusive

APP 3.5 requires Commonwealth agencies to only collect personal information by lawful and fair means.

1.5 Direct collection

APP 3.6 requires Commonwealth agencies to collect personal information about an individual from that person unless:

- (a) the individual consents to the agency collecting the information from a third party (APP 3.6(a)(i));
- (b) the agency is required or authorised by or under an Australian law to collect the information from a third party (APP 3.6(a)(ii)); or
- (c) it is unreasonable or impracticable to collect the information directly from the individual (APP 3.6(b)).

1.6 Collection transparency

APP 5 requires agencies, when collecting personal information about an individual, to take such steps as are reasonable in the circumstances to ensure that the individual is aware of a number of matters, which include the following:

- (a) the purposes for which the individual's personal information is being or has been collected;
- (b) whether the collection is authorised or required by law;
- (c) if the agency collects, or has collected, personal information from a third party, the fact that the agency so collects or has collected the information, and the circumstances of the collection; and
- (d) any third parties to whom it is the agency's usual practice to disclose that information.

2. Use and disclosure

2.1 Use and disclosure of personal information

APP 6 imposes limitations around the use and disclosure of personal information for a purpose other than the purpose for which the information was collected in the first place (i.e. a secondary purpose). There are also additional restrictions in relation to the use and disclosure of 'sensitive information'.

2.2 Use and disclosure of unique identifiers

APP 7 limits the collection and use of government-issued identifiers by private sector organisations. There are no limitations under the APPs in relation to the collection, use or disclosure of government-related identifiers by agencies.

2.3 Cross-border disclosures

APP 8 imposes additional restrictions in relation to the disclosure of personal information to an overseas recipient.

3. Data quality

APP 10 requires agencies to take reasonable steps (if any) to ensure that:

- (a) the information they collect is relevant to the purpose of collection, and is up to date and complete; and
- (b) the information is accurate prior to using or disclosing it.

4. Data security

APP 11 requires Commonwealth agencies to take reasonable security safeguards to ensure that the information they hold is protected against loss, unauthorised access, use, modification or disclosure, and against other misuse.

5. Access and correction

APP 12 requires agencies to provide individuals with access to their own personal information, unless an exception applies (for example, the agency is required or authorised to refuse access under another Commonwealth law that provides for access to documents, such as the *Freedom of Information Act 1982* (Cth)).

APP 13 requires agencies to take reasonable steps to correct any personal information they hold to ensure it is accurate, relevant, up to date, complete and not misleading.

Schedule 8 - Sources of information

Publicly available information

ABS

- *ABS Privacy Policy* (<http://www.abs.gov.au/websitedbs/D3310114.nsf/Home/Privacy+Policy>) (as at August 2017)

AGDET

- 2016-2017 National Partnership Agreement on Universal Access to Early Childhood Education (available at <https://www.education.gov.au/national-partnership-agreements>)

NSS

- *A Guide for Data Integration Projects Involving Commonwealth Data for Statistical and Research Purposes* (available at <https://statistical-data-integration.govspace.gov.au/topics/applying-the-separation-principle>)
- *Data Integration Involving Commonwealth Data for Statistical and Research Purposes: Governance and Institutional Arrangements* (6 October 2010)
- *High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes* (3 February 2010)

OAIC

- *Australian Privacy Principles Guidelines*
- *Community Attitudes to Privacy Survey 2017*
- *Guide to securing personal information: 'Reasonable steps' to protect personal information* (January 2015)
- *Guidelines on Data Matching in Australian Government Administration* (June 2014)
- *Information Policy agency resource 1: De-identification of data and information* (April 2014)
- *Privacy Business Resource 4: De-identification of data and information* (April 2014)
- *Privacy Impact Assessment Guide* (May 2014)

Documents provided by the ABS

- *ABS Data Integration Project Proposal: NPF VDC Project*, including the internal Privacy Impact Assessment prepared by the ABS (set out in Attachment 2)
- *ABS Data Integration Protocol*
- AESOC, Early Childhood Policy Group, *National Project Fund Bid: Vulnerable and Disadvantaged Children* (December 2016)
- AESOC, Early Childhood Policy Group, *National Project Fund Bid: Vulnerable and Disadvantaged Children – Summary Responses* (17 February 2017)
- *Census Data Enhancement project: An update, 2011 (cat. no. 2062.0)* (available at <http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/2062.0Main%20Features42011?open=document&tabname=Summary&prodno=2062.0&issue=2011&num=&view=>)
- *Letter to State Education Authorities* (template)
- *Responsible Use of ABS Microdata: User Guide (cat. no. 1406.0.55.003)*
- *VDC Input Data Items List*

Legislation & Legislative Instruments

- *A New Tax System (Family Assistance) (Administration) Act 1999* (Cth)
- *Census and Statistics Act 1905* (Cth)
- *Census and Statistics Regulation 2016* (Cth)
- *Freedom of Information Act 1982* (Cth)
- *Family Assistance (Public Interest Certificate Guidelines) Determination 2015* (Cth)
- *Privacy Act 1988* (Cth)
- *Social Security (Administration) Act 1999* (Cth)

Other Instruments

- *A New Tax System (Family Assistance) (Administration) Act 1999 – Public Interest Certificate*, made in 2017 by a delegate of the Secretary, AGDET under paragraph 168(1)(a) of the FAA Act
- *Social Security (Administration) Act 1999 – Protected Information Disclosure* instrument, made in 2017 by a delegate of the Secretary, DSS under subsection 202(2C) of the SSA Act

Schedule 9 – Glossary and acronyms

Term or Acronym	Meaning
ABS	Australian Bureau of Statistics
ADAU	Administrative Data Acquisition Unit, which is a unit within the ABS which undertakes the acquisition and management of electronic administrative data transfer by providers and transformation of the data for internal clients.
AEDC	Australian Early Development Census
AESOC	Australian Education Senior Officials Committee
AGDET	Australian Government Department of Education and Training
anonymisation	The process of converting name and address information into a non-identifying group code.
APPs	Australian Privacy Principles
APP Guidelines	The <i>Australian Privacy Principles Guidelines</i> (as at 1 April 2015), prepared by the Australian Information Commissioner under section 28 of the Privacy Act.
Catholic Education Office	Catholic Education Office of the Archdiocese of Canberra and Goulburn
CCMS	Childcare Management System
CCMS PIC	<i>A New Tax System (Family Assistance) (Administration) Act 1999 – Public Interest Certificate</i> , made in 2017 by a delegate of the Secretary, AGDET under paragraph 168(1)(a) of the FAA Act
Census and Statistics Act	<i>Census and Statistics Act 1905</i> (Cth)
Census and Statistics Regulation	<i>Census and Statistics Regulation 2016</i> (Cth)
COAG	Council of Australian Governments
CRN	Customer Reference Number (an identifier assigned by DHS to Centrelink customers)
DHS	Department of Human Services
DLC	The ABS Data Linkage Centre which undertakes data linkage projects, and builds organisational capability in data linkage to enhance the evidence base for decision making, research and discussion in government and the community.
DOB	Date of birth
DSS	Department of Social Services
DSS PID	<i>Social Security (Administration) Act 1999 – Protected Information Disclosure</i> , made in 2017 by a delegate of the Secretary, DSS under subsection 202(2C) of the SSA Act.
ECEC	Early Childhood Education and Care
ECEP	Early Childhood Education Programs
ETS	Education and Training Statistics
FAA Act	<i>A New Tax System (Family Assistance) (Administration) Act 1999</i> (Cth)
FOI Act	<i>Freedom of Information Act 1982</i> (Cth)

Term or Acronym	Meaning
FTB	Family tax benefit
Governance Arrangements	<i>Data Integration Involving Commonwealth Data for Statistical and Research Purposes: Governance and Institutional Arrangements</i> (6 October 2010)
IRAP	Information Security Registered Assessors Program
Mesh Block	The smallest geographical area defined by the ABS, identified in the 2011 Australian Statistical Geography Standard.
NAPLAN	National Assessment Program Literacy and Numeracy
NAPLAN Data Custodians	The state and territory education departments and agencies, and NAPLAN test administration authorities, identified in Schedule 4.
NECECC	National Early Childhood Education and Care Collection
NSS	National Statistical Service
OAIC	Office of the Australian Information Commissioner
personal information	As defined in subsection 6(1) of the Privacy Act.
PIC	Public Interest Certificate
PIC Guidelines	<i>Family Assistance (Public Interest Certificate Guidelines) Determination 2015</i> (Cth)
PID	Protected Information Disclosure
Privacy Act	<i>Privacy Act 1988</i> (Cth)
Privacy Policy	ABS Privacy Policy (published on the ABS website as at July 2017: http://www.abs.gov.au/websitedbs/D3310114.nsf/Home/Privacy+Policy).
SA1	Statistical Area Level 1 defined by the ABS, identified in the 2011 Australian Statistical Geography Standard.
SAA Act	<i>Social Security (Administration) Act 1999</i> (Cth)
SEIFA	Socio-Economic Indexes for Areas
SLK	Statistical Linkage Key, which is a code that enables two or more records relating to the same individual to be brought together.
SOM	Senior Officials Meeting
SSRI	Social Security and Related Information
TAA	Test Administration Authority
VDC	Vulnerable and disadvantaged children
2011 Population Census	2011 Census of Population and Housing