



2021 Census Administrative Data privacy impact assessment

For: Australian Bureau of Statistics

Date: March 2020

COMMERCIAL IN CONFIDENCE

**INFORMATION
INTEGRITY
SOLUTIONS**

managing the **privacy** of **individuals**
is **complex** and we can help you get
it **right**

Table of contents

1. Executive summary	4
1.1 IIS's overall view	4
1.2 Summary of recommendations	5
2. Introduction	7
2.1 Scope and methodology	7
2.2 Stakeholder engagement	7
2.3 Structure of the report	7
3. Applicable laws	8
3.1 The <i>Privacy Act 1988</i>	8
3.2 ABS's enabling legislation	8
3.3 Legislation regulating data custodian disclosure	9
3.4 Relationship to the Multi Agency Data Integration Project	9
4. About the Census admin data project	11
4.1 Meaning of 'administrative data'	11
4.2 Is admin data personal information?	12
4.3 Use case 1 – Maximising the Census response	12
4.4 Use case 2 – Improving the Census count	13
4.5 Use case 3 – Repairing the Census	14
4.6 Information flows and data separation	15
4.7 Governance and oversight	15
4.8 Policies governing data integration	16
4.9 Data sharing agreements	16
5. KEY RISK AREA 1 – Strategic considerations	17
5.1 ABS future directions on admin data use	17
5.2 Privacy implications of greater use of admin data	17
5.3 Community attitudes and social licence	19
5.4 Engaging with groups representing the interests of civil society	20
5.5 Threshold for pursuing use case 3	21
5.6 Impact of the Data Availability and Transparency Bill	22
6. KEY RISK AREA 2 – Transparency	23
6.1 Privacy policy (APP 1)	23
6.2 Privacy notice (APP 5)	23
6.3 Explaining admin data use to the community	24

6.4	'Admin data' as a concept	26
6.5	Use of admin data for enforcement.....	26
7.	KEY RISK AREA 3 – Electricity usage data	27
7.1	Electricity usage data and smart meters.....	27
7.2	Is electricity usage data personal information?	27
7.3	Privacy impact and stakeholder feedback	28
7.4	Data minimisation.....	29
8.	Other considerations	30
8.1	Sensitive information.....	30
8.2	Collection (APP 3).....	30
8.3	Use (APP 6)	31
8.4	Disclosure (APP 6).....	32
8.5	Security (APP 11).....	32
9.	Appendix A – Glossary	34
10.	Appendix B – Data variables involved	36
11.	Appendix C – Information flows and data separation.....	38
11.1	Collection.....	38
11.2	Use	38
11.3	Disclosure.....	39
11.4	Disposal.....	39
12.	Appendix D – APP compliance table.....	40
13.	Appendix E – PIA methodology.....	45
14.	Appendix F – Documents reviewed	46
15.	Appendix G – Stakeholder meetings held	48

1. Executive summary

The Australian Bureau of Statistics (ABS) contracted Information Integrity Solutions Pty Ltd (IIS) to conduct a Privacy Impact Assessment (PIA) on a proposal to use integrated administrative data (admin data) in the next Census. The project (henceforth, the Census admin data project) involves three use cases:

- **Maximising the Census response**
Using admin data to help ABS identify localities that require additional or specific forms of support to participate in the Census (for example, ABS may endeavour to provide more hard copy forms to an area with a high proportion of older Australians)
- **Improving the Census count**
Using admin data to better determine whether a house that did not return a Census form was occupied on Census night and using the data to choose more compatible 'donor houses'.
- **Repairing the Census**
In the event of a natural disaster or an across the board low response to the Census, using admin data to repair Census data.

ABS already collects admin data for various purposes, including in relation to the Multi Agency Data Integration Project (MADIP). The Census has used administrative data in the past to support the count and improve data quality, for instance using it to enumerate prisons and Australian Antarctic stations. The difference with this project is that the admin data will be integrated, meaning that a number of admin datasets will be brought together and merged. Moreover, although it operates under much of the same governance infrastructure as MADIP, it falls outside of the operation of MADIP. Unlike MADIP, the admin data is for internal use by ABS rather than external use by researchers.

IIS has conducted this PIA in conjunction with a consultation process run by ABS. It also engaged in a number of meetings with ABS staff to learn about the project. IIS would like to thank ABS for its assistance during the PIA process. ABS staff have been helpful and forthcoming throughout the drafting of the PIA and made every effort to assist IIS with its enquiries and provide information.

1.1 IIS's overall view

Generally, IIS finds that ABS has robust governance arrangements in place for its data integration activities. Those arrangements mean that the data processing proposed under the three use cases would already meet a high standard in terms of privacy best practice. Many of the Australian Privacy Principle (APP) compliance issues that might otherwise arise have already been addressed and rectified in earlier MADIP PIAs.

IIS also finds that there are many privacy enhancing measures in place to further minimise the privacy impact. Here, we are particularly referring to the data separation measures in place, along with processes in place to remove identifiers and work as much as possible with [unidentified data](#).

The issues that remain relate to fostering and maintaining the trust of the community. Assessments like this PIA offer an opportunity to apply a wider frame to the project and understand it in a broader organisational context. Failing to address this wider frame means that incremental change occurs

without a broader check. The problem with incremental change is that each increment seems reasonable on its own until the tipping point when the increments add up to a bigger risk than the organisation planned on taking.

IIS has, therefore, structured this report around three core themes, rather than addressing each APP in turn. Where relevant, IIS addresses particular APPs within those themes. Additional APP considerations are at section 8 and a summary of IIS's assessment against the APPs is at [Appendix D](#).

Two core areas identified by IIS are:

- **Strategic considerations**
This includes ensuring that ABS monitors cumulative privacy impact and involves groups representing the interests of civil society as much as possible in consultations about admin data use. ABS should also take a considered and strategic approach to pursuing use case 3. See section 5 and recommendations 1-4.
- **Transparency**
This includes updating privacy policies and notices (in line with APPs 1 and 5), presenting clear and cohesive information about admin data use and identifying additional avenues for communicating this information. See section 6 and recommendations 5-9.

A third core area explored in this report was ABS's proposed use of electricity usage data. ABS has since confirmed that it will not go ahead with use of dwelling-level electricity data in the 2021 Census. For transparency, IIS has retained its advice on this matter in [section 7](#):

- **Dwelling-level electricity usage data**
This includes making a strong case for any future use of this information and how the improvement to Census figures is large enough to justify the impact on privacy. See [section 7](#) and recommendation 10.

IIS also recommends that ABS explore options to reduce the amount of data collected if it decides to pursue use case 3 (see [section 8](#) and recommendation 11).

1.2 Summary of recommendations

A summary of recommendations appears below. Recommendations appear in full in the body of the report, in sections 5 to 8.

Recommendation 1 – Ensure that expansions to admin data use are reviewed at a sufficiently high level within the agency and are subject to scrutiny on cumulative privacy impact.

Recommendation 2 – Facilitate participation of groups representing the interests of civil society in consultations about admin data use and remove barriers to participation.

Recommendation 3 – Establish a threshold test for pursuing use case 3 in a case of localised low response.

Recommendation 4 – In a case of across-the-board low response, set a specific low response threshold in advance that makes clear what response rate is low enough to trigger use case 3.

Recommendation 5 – Update privacy policy to reflect changes associated with this project.

Recommendation 6 – Work with data custodians to meet APP 5 requirements

Recommendation 7 – Develop and deploy a communications plan and identify additional methods for informing individuals about admin data use.

Recommendation 8 – Use meaningful, plain-English terms in communications to individuals.

Recommendation 9 – Clarify that admin data will not be used for enforcement.

Recommendation 10 – Conduct and publish further in-depth analysis of the level of improvement to Census data the use of electricity data would bring about before proceeding.

Recommendation 11 – Explore options to reduce the amount of data collected in the event that ABS pursues use case 3.

2. Introduction

The Australian Bureau of Statistics (ABS) contracted Information Integrity Solutions Pty Ltd (IIS) to conduct a Privacy Impact Assessment (PIA) on a proposal to use integrated administrative data (admin data) in the next Census. The project (henceforth, the Census admin data project) involves three use cases, described in [section 4](#).

2.1 Scope and methodology

In carrying out a PIA, ABS asked IIS to:

- Identify privacy issues and risks associated with the Census admin data project – including matters of compliance with law and policy, as well as broader considerations such as stakeholder expectations and social licence.
- Make recommendations to mitigate or remove the privacy impacts.

Note that this PIA on the Census admin data project is separate to the 2021 Census PIA, which was conducted concurrently with this one.

IIS conducted the PIA in a series of phases, including planning, information gathering, consultation, analysis and drafting the report. More information about IIS's methodology for the PIA is at [Appendix E](#).

2.2 Stakeholder engagement

The ABS organised three two-hour roundtable sessions with external stakeholders in Canberra, Sydney and Melbourne (with people attending via video conference from Brisbane). During these sessions, ABS staff gave a presentation of the Census admin data project and the three proposed use cases. Stakeholders were able to offer feedback during the presentation and afterwards. They were also given the opportunity to provide further input by email afterwards.

IIS staff attended the sessions but did not participate in the discussion. The objective was to listen and hear the feedback of stakeholders. IIS has taken that feedback into account during the analysis and drafting of this PIA. Where relevant, we have quoted the views of stakeholders through the body of the report. For example, we have included in detail comments about electricity usage data as this was a topic of significant interest and concern for stakeholders. We also produced a short summary of the key issues raised and comments made in the roundtable sessions which is to be made available alongside this report.

A list of the stakeholders consulted is provided at [Appendix G](#).

2.3 Structure of the report

Information about applicable laws and a description of the three use cases are set out in sections 3 and 4. IIS's findings and analysis are set out in sections 5, 6, 7 and 8. These sections also contain IIS's recommendations in full (alongside the supporting analysis). The appendices hold descriptive and supporting information.

3. Applicable laws

This PIA is primarily concerned with the *Privacy Act 1988* (Cth) (the Privacy Act) and compliance with the Australian Privacy Principles (APPs). However, other legislation (outlined below) also facilitates the data sharing to occur under the project.

3.1 The *Privacy Act 1988*

The ABS is covered by the Privacy Act and its 13 APPs. The APPs set rules for the handling of personal information which the Act defines as any ‘information or an opinion about an identified individual or an individual who is reasonably identifiable’ (s 6(1)).

The APPs impose a range of privacy enhancing obligations on information handlers. This includes concepts such as data minimisation, purpose limitation and security. The APPs also give individuals certain rights and choices in relation to their personal information which individuals can pursue under the Privacy Act’s complaint-handling and enforcement provisions. IIS has assessed the information flows associated with this project against each of the APPs.

The Census admin data project particularly engages the following principles: APP 1 (Transparent handling of personal information), 3 (Collection of solicited personal information), 5 (Notice), and 6 (Use and disclosure). Other principles such as APP 11 (Security) are important but are not changing for this project – that is, existing security arrangements apply to data integration activities and those arrangements have been assessed in previous PIAs; this project, which is a data integration project, operates within those arrangements.

Section 8 addresses compliance with key APPs in more detail. APPs 1 and 5 are also discussed in section 6. In addition, [Appendix D](#) offers a summary of IIS’s findings against each principle.

3.2 ABS’s enabling legislation

The ABS is authorised to collect, compile, analyse, and publish statistics under the *Australian Bureau of Statistics Act 1975* and the *Census and Statistics Act 1905*. While ABS may publish statistical outputs, these must not be published or disseminated in a manner that is likely to enable the identification of a particular person or organisation.¹ Data collected by the ABS is also protected by strict secrecy provisions in the *Census and Statistics Act*.

ABS’s legislation thus ‘unlocks’ Privacy Act restrictions on collection of personal information contained in APP 3. APP 3 allows an agency to collect personal information if it ‘is reasonably necessary for, or directly related to, one or more of its functions or activities’ (APP 3.1). That said, ABS’s legislation does not create a carte blanche for data collection. The requirement that the personal information be ‘reasonably necessary’ establishes a data minimisation test which aims to reduce privacy impact through lessening the amount of information collection to only that which is reasonably necessary. We discuss this matter further in [section 8.2](#).

¹ *Census and Statistics Act 1905*, s 12(2).

3.3 Legislation regulating data custodian disclosure

The Privacy Act restricts disclosure of personal information. Generally, an entity may only disclose personal information for the primary purpose it was collected unless an exception in APP 6 allows disclosure for a secondary purpose. The point of this principle is to arrest scope creep – where personal information collected for one purpose is used for other additional purposes outside the expectations of the individual.

Despite the ABS's authorisation to collect personal information for its statistics activities, data custodians are still covered by APP 6 and must comply with its terms to disclose the data. In this case, data custodians will rely on the APP 6 exception that allows disclosure for a secondary purpose where the disclosure is required or authorised by or under an Australian law (APP 6.2(b)).

The Australian Taxation Office is able to disclose personal income tax information to the ABS under provisions in the *Taxation Administration Act 1953* and the *Tax Law Amendment (Confidentiality of Taxpayer Information) Act 2010*. Other agencies – the Department of Health, the Department of Human Services and the Department of Social Services – do not have an explicit authorisation in their legislation and must therefore rely on Public Interest Certificates (PICs) to be issued for this project.² PICs may be issued under certain pieces of legislation to enable specific and limited disclosures of information in the public interest.

Energy distributors disclosing electricity usage data to ABS for use case 2 would do so under a letter of exchange with ABS. ABS has since confirmed with IIS that it will not go ahead with use of dwelling-level electricity data but may still use electricity data aggregated to the meshblock level. This data is likely to fall outside the definition of personal information, which means the Privacy Act will not apply to this disclosure. This issue is discussed further in [section 7.2](#).

3.4 Relationship to the Multi Agency Data Integration Project

The [Multi Agency Data Integration Project](#) (MADIP) is intended to be a secure and enduring approach for combining data from across government domains over time. It involves a cooperative arrangement between ABS and a number of other agencies, under which the agencies share their data to support research and data reuse. ABS plays the role of 'accredited Integrating Agency' – this involves integrating data from multiple sources and making it available in [de-identified form](#) to authorised researchers.

ABS already collects and integrates admin data in conjunction with MADIP. The Census admin data project that this PIA assesses is different from MADIP because the collection and integration is for ABS internal purposes only – the integrated data will not be made available to researchers. That said, this project benefits from much of the governance infrastructure that supports MADIP activities. This includes the MADIP Operating Model and the requirement to have a data integration plan in place for

² PICs will need to be issued under s 130 of the *Health Insurance Act 1973*, s 135A of the *National Health Act 1953*, s 208 of the *Social Security (Administration) Act 1999*; s 168 of the *A New Tax System (Family Assistance) (Administration) Act 1999*; s 128 of the *Paid Parental Leave Act 2010*; and s 355 of the *Student Assistance Act 1973*.

data integration projects. ABS's data integration activities for MADIP have also been the subject of two recent PIAs which have further identified and addressed privacy compliance issues associated with data sharing and integration.

4. About the Census admin data project

The Census admin data project involves use of ‘administrative data’ to improve Census data. It comprises three use cases which are each outlined below.

4.1 Meaning of ‘administrative data’

On its website ABS states that ‘administrative data refers to information maintained by governments and other entities that is made available to the ABS for statistical purposes. It includes data collected for registrations, transactions and record keeping, usually during the delivery of a service.’³

The types of datasets ABS intends to use for the Census admin data project are outlined in the table below.

Dataset	Custodian	Description
Person linkage spine	MADIP Board	A concordance or ‘map’ of the links between the core MADIP spine datasets that creates a combined MADIP population and is ‘scoped’ to a point in time using information from MADIP datasets.
Medicare Consumer Directory (MCD)	Department of Health Services Australia	Demographic information on persons enrolled with Medicare.
Personal Income Tax Client Register (PIT)	Australian Taxation Office	Demographic information on individuals who require a tax file number to interact with government, business, financial, educational and other community institutions.
Social Security and Related Information (SSRI)	Department of Social Services	Characteristics of recipients of Government payments such as Age Pension, Newstart Allowance, and Family Tax Benefit.
Electricity usage data	Electricity distributors (various)	Information on the connection status and volume of usage of electricity to residential electricity meters within a specified period of time.

Examples of analytic variables ABS wishes to use from the datasets outlined in the table above is further explained in sections 4.3 and 4.4 below. For further information about the data variables involved, see [Appendix B](#).

³ ABS, [Administrative Data Research for the 2021 Census](#), webpage, accessed 29 February 2020.

4.2 Is admin data personal information?

Admin data can include personal information such as name, address, and date of birth (linkage variables), along with other information about the individual (analytical variables). The linkage variables are stored separately from the analytical variables in accordance with the separation principle. The ABS explained to IIS that linkage variables were used to link the datasets to the central linkage infrastructure (known as the Person Linkage Spine). For this project, the main linkage variable to be used is address.

The analytical variables may, in some circumstances, be considered personal information even when they are separated from direct identifiers as they may enable the re-identification of an individual (for example, through the combination of data items). Regardless of whether the data contains personal information or not, ABS advised IIS that it treats the data as if it were personal information (with all the attendant privacy requirements and security obligations).

Despite this approach (of treating the data as personal information), there is, of course, still significant value in minimising the identifiability of the data as much as possible. For the Census admin data project, all data will be un-identified for internal ABS analytic purposes. This necessarily minimises the privacy impact on individuals and reduces risks in the event of a breach (accidental or intentional).⁴ Further information about the data separation and integration process is outlined in [Appendix C](#). Further information about data security is outlined in [section 8.5](#).

In addition to sourcing data from public sector agencies, ABS also proposes to collect electricity usage data from energy distributors to help determine dwelling occupancy on Census night. The data will be at the household level and will not include any 'retail information' about the customer and their billing details. See [section 7](#) for a discussion of the privacy considerations for this data.

4.3 Use case 1 – Maximising the Census response

Use case 1 involves planning the enumeration strategies before Census collection. ABS will use admin data to help determine areas for special approach, field staff needed and how they are deployed, and follow up efforts.

Some sub-population groups have lower response rates to the Census. For example, young people and Aboriginal and Torres Strait Islander people both have lower response rates from Census to Census. To encourage a higher response rate, ABS deploys strategies to encourage Census participation or raise community awareness. ABS also makes special approaches to other sub-populations, such as older people who may prefer paper forms, and groups with cultural or language barriers who may need extra support. In all cases, this necessitates ABS understanding where such populations are concentrated so that it can tailor its approach in those areas.

⁴ ABS also has an Incident Response Procedure in place to enable a swift response to a data breach.

In the past, ABS has used past Census data to locate (to the area level) where sub-populations are concentrated. However, past Census data is already five years old at this point and potentially out-of-date. Admin data offers more up-to-date information.

The ABS proposes to collect age, Aboriginal and Torres Strait Islander status, and country of birth data from agencies that hold it. Data analysts within the ABS would aggregate the data to a small area level (with each area block containing 30-60 dwellings) and pass the aggregated data to the responsible Census Team. That Census Team is then equipped to allocate resources and tailor enumeration strategies accordingly.

The type of data involved in use case 1 is set out in the table below:

Analytical variable needed for analysis	Datasets used	Level of identification
Demographic characteristics (age, sex)	MCD, PIT, SSRI	Aggregated to geographic area (meshblock level – 30-60 dwellings)
Indigenous status, country of birth alongside 2016 Census outcomes	SSRI	Aggregated to geographic area (meshblock level – 30-60 dwellings)

4.4 Use case 2 – Improving the Census count

Admin data for use case 2 would be used after Census night, during the data processing phase (a period of up to 18 months after the Census).

Use case 2 would aim to address two issues:

- Difficulties in determining dwelling occupancy (that is, ABS’s ability to determine whether a dwelling was empty on Census night).
- Imprecise use of ‘donor houses’ (in cases where ABS finds that a dwelling was probably occupied on Census night, despite no one responding, and ABS fills in the gap with ‘dummy data’ by borrowing and imputing data from a similar nearby household that did respond – a donor).

With regard to the first bullet, analysis after the 2016 Census revealed that there was an increase in the number of empty dwellings which were incorrectly determined to be occupied as compared with the 2011 Census. Mistakes of this type can affect the accuracy of population counts. Therefore, ABS proposes to use admin data to calculate, with more accuracy, the likelihood of dwelling occupancy. It would do this by combining data from the Australian Taxation Office and the Department of Social Services (showing currency of address information) to create a score indicating occupancy likelihood.

As indicated in the second bullet, ABS assigns ‘dummy’ information (age, sex and marital status) to households that did not respond. However, imputing dummy data can lead to inaccuracies as donor houses tend to over-represent older Australians (who are more likely to have responded to the Census).

ABS proposed to use admin data to allow it to choose donor houses that are closer to the occupied houses they stand in for. To do this, ABS would collect admin data indicating the number of residents at a given address by age and sex and then match the house with a donor household with a similar make-up.

The type of data involved in use case 2 is set out in the table below.

Analytical variable needed for analysis	Datasets	Level of identification
<p>Count of males in a dwelling (aggregate at dwelling-level, derived item)</p> <p>Count of females in a dwelling (aggregate at dwelling-level, derived item)</p> <p>Count of persons by age groups (aggregate at dwelling-level, derived item)</p>	MCD, PIT, SSRI	Only dwelling system ID is provided to the ABS analysts. No address or other personal information is provided.
Occupancy probability for a dwelling – derived item not raw data (modelled using a range of administrative data)	Derived using PIT, SSRI, electricity usage data.	Only dwelling system ID is provided to the analysts. No address or other personal information is provided.

Following initial advice from IIS, ABS has opted not to proceed with using dwelling-level electricity usage data in the Census at this time. This means that use case 2 is unlikely to proceed in the form outlined above, though ABS indicated that it may revisit this matter for future Censuses. Instead, ABS is exploring options to use electricity data aggregated to the meshblock level. IIS finds that this data is unlikely to meet the definition of personal information which means the Privacy Act will not apply. For transparency, IIS has retained its advice on collection and use of electricity data in [section 7](#) of this report.

4.5 Use case 3 – Repairing the Census

Use case 3 involving repair of Census data with admin data would only be pursued if necessary. It is contingent on a low Census response rate, either across the board or in a certain area (due to a natural disaster, for example). Low response rates occurred in Canada in 2016 (due to a bushfire in the Fort McMurray region which severely limited Census response rates in that area) and in New Zealand (which experienced a low response rate across the board and successfully used admin data to fill some of the gaps).

ABS has indicated that Census repair may involve an extension of use case 2 whereby ABS uses admin data to help guide its choice of 'donor houses' to stand in for missing households. Alternatively, it may involve imputing admin data directly into the Census dataset. The approach taken will depend on the circumstances and ABS has committed to conducting an additional PIA in the event that it pursues use case 3.

4.6 Information flows and data separation

ABS collects the admin datasets outlined above (and in more detail in [Appendix B](#)) and then sends the data through a multi-stage process to anonymise and merge the datasets ready for use for the use cases outlined above. A detailed description of information flows and data separation is at [Appendix C](#).

4.7 Governance and oversight

The 2021 Census Program is the project owner for the Census admin data project. This project is subject to the Census governance arrangements, with oversight at various levels, including by the:

- **2021 Census Executive Board** oversees the strategic direction of the 2021 Census Program and makes decisions about policy, priorities and direction for the 2021 Census Program.
- **Senior Responsible Officer** is responsible for major decisions, operations and outcomes for the 2021 Census Program.
- **2021 Census Data Operations Branch Program Manager** provides oversight of the 2021 Census admin data project.
- **2021 Census Delivery Committee** is the primary forum to advise, monitor and review the 2021 Census Program and reports to the 2021 Census Executive Board.

Additionally, the ABS has well-established oversight infrastructure for data integration projects. The operational oversight of the Census admin data project will be by the General Manager (SES Band 2 officer) that heads up the Statistical Data Integration Division. This is standard practice for data integration projects. Some projects also receive oversight higher up the line, depending on factors such as strategic risk. Higher level oversight rests with the:

- **ABS Executive Board** for enterprise strategic risk oversight
- **ABS Data Innovations and Statistical Strategy Committee** for data integration program strategic risk oversight
- **ABS Data Integration Program Board** for data integration program risk management.

Although the Census admin data project falls outside of MADIP, ABS has leveraged MADIP frameworks and presented the proposal to the MADIP Board – an external group of MADIP data custodians – to test project arrangements.

4.8 Policies governing data integration

The MADIP Operating Model outlines technical and procedural requirements for data integration projects. It also details standard processes used throughout the end-to-end journey for data including in relation to project approval, data linking, assembling linked data extracts, data access, confidentiality, communication, data retention and destruction, and breach management.

ABS uses data integration plans as the key governance document applying to data integration projects. They help ABS meet its responsibilities as an [accredited integrating authority](#). They also offer 'a check' that the project in question is of benefit to the public, in line with the [data integration high level principles](#). Plans use a standard template that includes information on:

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

4.9 Data sharing agreements

ABS has a data sharing agreement with each data custodian which meets relevant legal requirements. However, it does not have a single standardised agreement for these arrangements. Some arrangements are formalised via an MOU, while for others an exchange of emails is sufficient. IIS considered whether a standard approach might be desirable – ensuring, for example, that certain conditions and requirements associated with the sharing were consistently articulated and applied with each data custodian. If it were not for the existence of data integration plans – which set out much of this detail in a standard way – then the current approach might be deficient. The data integration plans, however, ensure key information is articulated including the datasets to be used, linkage and integration approaches, limits on data use and risk management. For that reason, IIS has decided not to make a formal recommendation on this point. However, we suggest ABS consider whether a more standardised approach is needed for agreements with custodians. We also note that the Data Availability and Transparency legislation is likely to introduce formal requirements in relation to any such agreements and so ABS may wish to wait to align with those requirements.

5. KEY RISK AREA 1 – Strategic considerations

ABS asked IIS to assess three use cases within the Census admin data project. Before we address particular privacy risks associated with those use cases, IIS would like to raise some higher-level strategic considerations. Assessments like this PIA offer an opportunity to apply a wider frame to the project and understand it in a broader organisational context. Failing to address this wider frame means that incremental change occurs without a broader check. The problem with incremental change is that each increment seems reasonable on its own until the tipping point when the increments add up to a bigger risk than the organisation planned on taking. IIS believes small expansions to admin data use could add up in this way, potentially exposing ABS to a risk outside of ABS's risk appetite.

5.1 ABS future directions on admin data use

ABS has indicated on its website that it is investigating options to add new information to the Census using admin data.⁵ ABS is reportedly also planning to research what Census content is already available from admin data and whether ABS could add income data to the Census using admin data.⁶

There was some concern about this amongst stakeholders consulted by ABS and IIS. One stakeholder expressed concern about the shift from 'Australians participating in the Census and doing their civic duty' to 'ABS will take your data.' Others wondered what was to stop ABS from using admin data for the whole Census. This Census admin data project is a long way from that scenario. However, it could be seen as a small step in that direction that may justify subsequent admin data use.

The movement to foster greater sharing of public sector data – as evidenced in the government's development of the Data Availability and Transparency Bill (see [section 5.6](#) below) – demonstrates that ABS is moving with the times; the Census admin data project is just one small part of a bigger story. As an accredited integrating authority and a key participant in MADIP, ABS is an important participant in efforts to unlock the value of public sector data. By implication, this means ABS is also likely to be influential in wider discussions and decisions about privacy and where the balance lies. Others may take their cue from ABS and the decisions it makes.

5.2 Privacy implications of greater use of admin data

The Census admin data project involves fairly restricted use of admin data. For use cases 1 and 2, admin data will not be imputed directly into Census figures. However, use case 3 does contemplate direct use of admin data in the Census. Wider privacy implications to consider in pursuing this direction include the following:

- **Public sector data custodians hold personal information of individuals who generally do not have a choice about interacting with the custodians**

⁵ ABS, [Administrative data research for the 2021 Census](#), webpage, accessed 29 February 2020.

⁶ ABS, [Administrative data research for the 2021 Census](#), webpage, accessed 29 February 2020.

KEY RISK AREA 1 – Strategic considerations

In other words, individuals already have limited choice about those agencies holding their information; this puts the onus on public sector bodies to hold themselves to a higher ethical standard than other sectors when it comes to data use and reuse. Reuse, in isolation, is not inherently bad and, indeed can support social and economic outcomes. However, broad-scale reuse may unfairly interfere with individuals' rights to live lives free from arbitrary intrusion or monitoring.

- **Taking information rather than asking for it reduces individual control and choice**

There is an ethical question that ABS must address (and indeed is probably in the process of addressing) about where the line lies between helping to 'reduce respondent burden' through collecting data from third parties and imposing its data collection agenda on the population (notwithstanding good intentions and the potential societal benefits of such data collection). While there is a wider societal return on statistics generated by the Census and other surveys, that return should not automatically override the public interest in protecting privacy. A careful 'weighing' must occur to determine the right balance on a case by case basis.

- **Greater use of admin data risks enabling a detailed (and potentially intrusive) view of individuals and their personal lives**

In the past, practical barriers would have limited the ABS's ability to collect and use large amounts of third-party data. However, the advent of digital technology and data analytics has opened up new frontiers in data collection and reuse. ABS now faces a situation where it may feasibly collect and use ever greater and more varied datasets about the population from a growing number of data custodians. These conditions exert considerable pressure on the ABS to continue to expand collection and use of third-party data. Against this backdrop, ABS must take care to consider the growing impact on individual privacy and the build-up of detailed information about individuals.

- **Greater use of admin data risks scope creep**

Admin data was collected for a particular purpose (allowing an individual to receive healthcare, for example, or lodge a tax return) and the subjects of the data may not have expected it to be collected and used by ABS. The government has authorised secondary use of such data via provisions in ABS's enabling legislation and other laws, including the Privacy Act (which enables secondary use in certain prescribed circumstances). Advances in technology, however, create conditions for data collection and reuse far beyond what was possible even ten years ago, or indeed when the legislation was enacted.

Of course, there are mitigating factors. For example, the privacy impact may be significantly reduced through use of data separation and removal or encoding of identifiers, and scope creep may be checked by governance in the form of data integration plans, strict [disposal schedules](#) and so on. The point is that, as with other similar projects (including MADIP), ABS must navigate some complex questions about privacy and ethics– not just 'do we have enough privacy safeguards in place to protect the data?' but 'should we collect this data at all?' In doing so, ABS must determine where the line lies between reasonable third-party data use and overreach. And it must be prepared to turn down opportunities where data is available and of interest but the cost – in terms of privacy intrusion, in terms of interference of 'the state' in the private lives of its individuals – is too great.

PIAs like this one help to broach such questions on a case by case basis. However, continued wider consideration and leadership at a strategic level will also be crucial.

Recommendation 1 – Ensure that expansions to admin data use in connection with the Census are reviewed at a sufficiently high level within the agency and are subject to scrutiny on cumulative privacy impact

ABS should ensure that it has measures in place to monitor (at a senior level) the accumulating privacy impact of incremental expansions to collection and use of third-party data in connection with the Census. ABS should allocate responsibility for this monitoring to an appropriate group within the agency.

The group should contribute to guiding the Census Division's strategic direction on increased admin data use.

The activities of the group should operate in alignment with, and advance the objectives of, the *Building Trust in the ABS and Our Data Use Strategy* and give particular regard to the Strategy's imperative: 'In everything we do, consider whether it builds or reduces trust'.

5.3 Community attitudes and social licence

The Productivity Commission explored the question of community support for data sharing and reuse in some detail in Chapter 3 of its 2017 *Data availability and use* inquiry report. It noted that the community generally does not view information sharing between departments as a major threat to privacy.⁷ The Office of the Australian Information Commissioner's (OAIC) 2017 community attitudes to privacy survey also indicated that government departments were the third most trusted type of entity (when it came to their handling of personal information).⁸ Indeed, anecdotal evidence suggests that most people expect that different parts of government share data; overseas studies show that people overestimate the extent of information sharing that is already occurring within government.⁹

Individuals would, however like to maintain a level of control over their information; they expect governments to share their data with their consent, only when strictly necessary, and to be transparent about their data handling processes.¹⁰ The Productivity Commission concluded that the onus is on government to communicate the benefits of data sharing effectively.¹¹

Trust can be fragile, and may be particularly so due to the issues associated with the 2016 Census. Moreover, community attitudes can change quickly, as evidenced by the My Health Record and the

⁷ See Productivity Commission, *Data availability and use: Inquiry report*, no. 82, 31 March 2017, p 123.

⁸ 58% of respondents said they trusted state and federal government departments, see OAIC, [Australian Community Attitudes to Privacy Survey 2017 Report](#), section 1.0.

⁹ See Productivity Commission, *Data availability and use: Inquiry report*, no. 82, 31 March 2017, p 123.

¹⁰ See Productivity Commission, *Data availability and use: Inquiry report*, no. 82, 31 March 2017, p 123.

¹¹ See Productivity Commission, *Data availability and use: Inquiry report*, no. 82, 31 March 2017, p 123.

transition from opt-in registration to opt-out. In the words of one stakeholder: ‘Trust is important. If you don’t get the basics right, you lose trust for any additional projects.’

Consultation with privacy advocates, academics and civil society revealed a level of discomfort about the breadth of admin data collection and use. Stakeholders specifically wanted to see evidence that the ‘return’, in the form of better occupancy determinations, for example, was great enough to justify the privacy impact. Stakeholders were particularly concerned by use case 2 and ABS’s proposal to use electricity usage data. This matter is discussed further in [section 7](#).

IIS finds that the major consideration for the Census admin data project is social licence. ABS needs to do as much as possible to foster it (including through greater transparency, see [section 6](#)). In some areas, despite being in compliance with privacy law, ABS may need to heed deficits in social licence and rethink or further justify proposed admin data use.

ABS is developing a *Building Trust in the ABS and Our Data Use Strategy*. IIS believes that a strategy of this nature is timely and worthwhile to guide ABS’s activities in this area.

5.4 Engaging with groups representing the interests of civil society

The consultation for this PIA comes after consultations for the MADIP PIAs and for the 2021 Census PIA. During consultation planning, there was some discussion between ABS and IIS about risks of ‘consultation fatigue’. ABS also heard from one civil society group that it did not have capacity to participate. It pointed out that, ‘the Australian Government is undertaking more and more data and technology work and expects frequent and deep engagement from community and consumer advocates but we are simply not resourced to participate properly, making for unbalanced policy-making.’

Now, more than ever, it is critical for ABS and other government agencies to ensure they receive input from groups representing civil society. As data processing becomes more complex, seamless and invisible to ordinary people, those people risk being marginalised in decisions about their own data. ABS may need to take a different or more supportive approach to engagement with civil society. This may be an issue ABS also considers in connection with its *Building Trust in the ABS and Our Data Use Strategy* and, in particular, its intention to ‘engage meaningfully’.

Recommendation 2 - Facilitate participation of groups representing the interests of civil society in consultations about data sharing and admin data use

ABS should explore and implement strategies to facilitate participation of groups representing the interests of civil society in consultations about data sharing and admin data use. Such groups may include civil liberties groups, consumer advocates, community groups, privacy advocates, academics and others. Strategies should aim to remove barriers to participation (noting that barriers often include resourcing constraints and other practical obstacles to participation).

5.5 Threshold for pursuing use case 3

Use case 3 potentially creates a greater privacy impact for two reasons. First, unlike the other use cases (which use area-level and address-level data) it may involve use of person-level data. Second, unlike the other use cases, it may involve direct imputation of admin data into the Census. That said, ABS has indicated that Census repair may involve an extension of use case 2 rather than direct imputation. It depends on the circumstances and the nature of the damage to the data. ABS has also committed to conducting another PIA (and presumably another Data Integration Plan) if it decides to pursue use case 3. IIS supports this approach.

IIS has raised use case 3 in this section about strategic considerations because pursuing use case 3 will be a strategic decision for ABS. In deciding whether to use admin data to repair the Census, ABS will need to consider a range of factors, including community expectations and what is reasonable in the circumstances.

Conducting a PIA on any proposal to deploy use case 3 will help to counteract the perception of scope creep, though it should be noted that by the time a PIA is commissioned, generally an agency has already decided to proceed.

We believe there should be a step before this where ABS formally assesses whether to proceed. IIS assumes ABS would do this anyway. What we are proposing, however, is that ABS sets the terms of that assessment beforehand to ensure transparency about its intentions and the factors it will take into account when deciding whether or not to proceed. In so doing, ABS should allow for a degree of flexibility to reflect the inherent uncertainty associated with conditions that may cause a low response. This may mean that a threshold test for use case 3 *guides* rather than *prescribes* ABS's decision-making process.

Recommendation 3 - Establish a threshold test for pursuing use case 3 in a case of localised low response

ABS should establish a threshold test to use to determine whether to pursue use case 3 in the case of a localised low response to the Census. The test should articulate variables that must be considered in decision-making. For example, this could include:

- Size of area affected
- Number of individuals affected
- Nature of the event that caused the low response (for example, the type of natural disaster) and whether this impacts on the feasibility of Census repair approaches
- Whether other methods, such as delayed survey response intake for a given area are feasible or desirable
- Specific impact of the low response on Census as a whole
- Specific impact of the low response on the area in question.

The combination of these (and other specified) factors should lead to an overall finding of 'significant damage to dataset' for use case 3 to be pursued. ABS could also quantify what 'significant damage' means in real terms to help guide its decision-making.

Recommendation 4 - Set a low response rate trigger point for use case 3 in a case of 'across-the-board' or widespread low response

ABS should determine a low response rate trigger point before the 2021 Census which makes it clear to the community upfront what ABS believes is a low enough across-the-board response rate to set in motion use case 3. It may wish to consider the experience of overseas jurisdictions like New Zealand in deciding where the threshold should lie.

5.6 Impact of the Data Availability and Transparency Bill

The ABS is engaging in the development of the Office of the National Data Commissioner's (ONDC) data sharing scheme (and draft bill) and made a submission to the ONDC's discussion paper in October last year. The data sharing scheme aims to facilitate and encourage greater sharing of public sector data. ABS also partnered with the ONDC to develop the *Best practice guide to applying data sharing principles* – principles which are modelled on the Five Safes framework.¹² Depending on when the bill is passed and what it contains, the scheme may enable ABS to receive data from other agencies without the need for a public interest certificate. However, it may also introduce new and different obligations, such as requirements to seek consent.

Until we know further details, it is difficult to gauge the impact it will have on this project. IIS understands that ONDC will release an exposure draft of the bill for consultation before the middle of the year.¹³ We suggest that ABS ensure that any changes it makes to align with the ONDC's data sharing scheme (once it is operational) do not lessen the privacy protections it already has in place for the Census admin data project (or put in place, in response to this PIA). If the change of approach is significant, an update to this PIA may be necessary.

¹² See ONDC, *Data sharing and release legislative reforms – Discussion paper*, September 2019, p 8.

¹³ See ONDC, *Data sharing and release legislative reforms – Discussion paper*, September 2019, p 53.

6. KEY RISK AREA 2 – Transparency

6.1 Privacy policy (APP 1)

APP 1 requires entities to have a privacy policy that covers certain (specified) matters and make the policy available to the public. ABS has more than one privacy policy – currently there is an ABS wide policy, a 2016 Census policy and a MADIP policy. The data handling under this project would be covered under the ABS privacy policy.

It is up to ABS how it decides to present its privacy policy (whether as one document or multiple). APP 1 does not specify the form a privacy policy must take, only that it be ‘clearly expressed’ and ‘up-to-date’. The OAIC does advise though that ‘for a large APP entity where distinct business units handle personal information differently, it may be appropriate for the entity to have a set of policies to cover the different types of personal information handled or different information handling practices.’¹⁴

Some stakeholders raised concerns about retention periods for integrated admin data and said that their concerns would be reduced if this information was clearly stated in a public facing document like the privacy policy. See recommendation 5.

Recommendation 5 – Update privacy policy to reflect changes associated with this project

ABS should update its privacy policy to reflect data handling changes associated with this project. In doing so, it should give regard to the OAIC’s [Guide to developing an APP privacy policy](#). As a matter of best practice, the policy should state the data retention arrangements for integrated admin data.

6.2 Privacy notice (APP 5)

APP 5 requires an entity to take reasonable steps either to notify the individual of certain matters or to ensure the individual is aware of those matters.¹⁵ This obligation applies even for personal information collected indirectly, which is to say from someone other than the individual. Therefore, APP 5 applies to ABS’s collection of admin data for this project. Stakeholders noted the importance of raising awareness of the Census admin data project with individuals, noting that ‘people didn’t know this data was going to be sent to ABS’ when they engaged with the data custodians.

ABS has already considered similar issues in its MADIP PIAs. In the most recent of those PIAs, the 2019 MADIP PIA update, ABS noted that it cannot update the collection notices of data custodians and that it was not reasonable for the ABS to directly notify, such as through a letter, each individual

¹⁴ OAIC, *APP Guidelines*, paragraph 1.11.

¹⁵ OAIC, *APP Guidelines*, Part 5.

represented in a dataset from one of those custodians.¹⁶ Instead, the ABS relies on the collection notices of data custodians that share data with ABS, other steps those entities may take to notify individuals, and other steps the ABS takes to build awareness of the collection and use of personal information in MADIP.¹⁷

In that PIA, ABS committed to increasing its own efforts regarding transparency and encouraging entities responsible for collection notices to update notices or otherwise make individuals aware of data use.¹⁸ To comply with APP 5 for this project, ABS will need to do something similar. Given that this appears to be an evolving space (potentially requiring further APP 5 notice updates in the future), ABS may want to explore options for data custodian privacy notices to all include the same link to a webpage on ABS's website that it updates as needed for this purpose. The webpage could offer the APP 5 information that ABS would otherwise offer in a notice direct to the individual.

Recommendation 6 – Work with data custodians to meet APP 5 requirements

ABS should, to the extent possible, work with data custodians to ensure privacy notices are fit for purpose in light of changes to occur under this project. To enable further flexibility in the future, ABS should explore options for data custodian privacy notices to all include (the same) link to an ABS webpage dedicated to explaining its admin data use (including in relation to this project) and other APP 5 matters. This would give ABS more control over APP 5 updates.

6.3 Explaining admin data use to the community

In its inquiry the Productivity Commission pointed out (and IIS concurs) that 'All development of data practice — whether in the private sector or public sector — must take the creation and preservation of understanding and trust as its first consideration.'¹⁹ ABS has already published information about admin data use. IIS's review of that information on ABS's website found that ABS has taken genuine steps to be transparent about its data collection and integration activities. There were webpages explaining ABS's intention to use admin data in the 2021 Census. There was also a large number of pages on the associated topic of MADIP and ABS's activities as an [Accredited Integrating Authority](#).

Feedback from stakeholders during consultations for this PIA emphasised the need for transparency about admin data use. Data custodians pointed out aspects of the project that would need careful explanation to allay concerns, including the meaning of 'administrative data' (discussed further below in [section 6.4](#)). They also noted that people were likely to ask: 'How is this new? Is this personal information? How will access to the data be managed?' and that it was important to be able to pre-empt those questions. Data custodians also spoke about the importance of giving a wider view of ABS's admin data use – that information about these specific use cases was important but there also

¹⁶ See ABS, *MADIP PIA Update*, 2019, p 35.

¹⁷ See ABS, *MADIP PIA Update*, 2019, p 35

¹⁸ See ABS, *MADIP PIA Update*, 2019, recommendation R2.

¹⁹ See Productivity Commission, *Data availability and use: Inquiry report*, no. 82, 31 March 2017, p 123.

needed to be more general information about admin data use. According to one stakeholder, it would be helpful to explain to the community ‘this is the ABS’s policy on the use of admin data generally.’ Another stakeholder said that it was also important to be transparent about risks, including risks of re-identification, and what measures ABS has in place to manage risks.

For some, existing information on the website raised more questions than it answered. One stakeholder pointed out that webpages about admin data use referred to collection of *de-identified* Medicare and SSRI data which to them ‘seemed highly unlikely.’²⁰ The webpages also failed to explain whether admin data would be used for enforcement (for example, to levy fines for failure to turn in a Census form). According to this stakeholder: ‘It sounds like it was written by someone trying to get away with something. It would be better to be up front and answer these questions.’ Other stakeholders thought that concerns they had with the project could be reduced through clear communication with the community about benefits and safeguards in place to minimise privacy impacts.

Getting the communications for this project right is important and challenging – ‘important’ because ABS needs to ‘bring the community along’ with it on admin data use and build trust through transparency; ‘challenging’ because both the Census admin data project itself and its relationship with MADIP are complex and multi-faceted. More information is not always better, particularly if the amount of explanatory information is overwhelming for an ordinary member of the public with no special knowledge of ABS’s functions and activities. It may be that presenting information in layers of detail or in easy-to-digest formats such as infographics or video will help.

ABS should also ensure that information on the Census admin data project and data integration is presented clearly and cohesively. Currently there are many webpages on data integration that appear to have been created and added to over time. The amount and complexity of the information could be challenging for some users. ABS acknowledges this and is carrying out a stocktake of its data integration information and assess whether the information needs to be restructured or revised to continue to present information as clearly and coherently as possible.

It goes without saying that the ABS’s website offers a powerful communications vehicle. However, it relies on individuals actively seeking out information. Such individuals will generally have a higher than average level of confidence and engagement with the subject area. One stakeholder said that publishing information on the website was a minimum: ‘Most people won’t be going to the ABS website – it’s not enough for something like this.’ ABS may need to use other channels to push information about admin data use, sharing and integration more generally, to ordinary Australians.

Recommendation 7 – Develop and deploy a communications plan and identify additional methods for informing individuals about admin data use

ABS should develop and deploy a communications plan (or extend its existing communications plan) for admin data use in the Census. The objective of the plan should be to enable ABS to take a considered and cohesive approach to raising awareness about, and communicating the details of, ABS’s admin data integration activities. This should include consolidating and making simpler

²⁰ Following this feedback, ABS advised that the use of the term ‘de-identified’ on the webpages in question was incorrect and that it would update web material.

the existing information on ABS's website. In developing the plan, ABS should also determine what other outreach it needs to carry out in addition to providing information online to proactively push information to individuals.

6.4 'Admin data' as a concept

IIS finds that the term 'administrative data' may not be meaningful to ordinary people seeking further information about the project. One stakeholder also observed that the term may not accurately reflect the inclusion of electricity usage data. ABS should consider ways to clarify this, including by better tailoring information to different audiences. For example, messaging to individuals may be clearer if it refers to collecting and using 'information about you' or 'information about you held by other government agencies,' rather than 'administrative data'.²¹ Stakeholders also suggested making clear what information was not included in the Census admin data project.

Recommendation 8 - Use meaningful, plain-English terms in communications to individuals

This may require ABS to consider an alternative to the term 'administrative data' for communications targeting individuals. If ABS uses terms like 'de-identified', 'un-identified' or 'anonymised' it should make sure it explains what they mean.

6.5 Use of admin data for enforcement

Numerous stakeholders from a range of sectors expressed concern that ABS might use admin data (particularly in connection to use case 2) to assist with imposing fines on individuals that fail to return a Census form. ABS assured stakeholders that it would not be using the data for that purpose. If that is the case, it should make this clear in its communications on the project and in relevant internal guidelines or procedures governing the project.

Recommendation 9 – Clarify that admin data will not be used for enforcement

ABS should ensure that its communications (and relevant internal governance) regarding admin data use make clear that admin data will not be used to identify and fine individuals who did not return a Census form or for any other enforcement activity.

²¹ See other jurisdictions guidance – for example, Stats New Zealand refers to '[real data about real people](#).'

7. KEY RISK AREA 3 – Electricity usage data

ABS's proposal use electricity usage data is markedly different from the other admin data use. The other data in this project comes from public sector data custodians, whereas electricity usage data will come from a mix of public and private sector energy distributors.

During the course of conducting this PIA, IIS submitted its initial findings and recommendations to ABS, including on ABS's proposed use of electricity data. In response to those preliminary findings, ABS decided not to proceed with use of dwelling-level electricity data in conjunction with use case 2. ABS indicated to IIS that it may instead use electricity data in aggregate form (for example, at the meshblock level) but was still exploring the feasibility of this. ABS also suggested that it may revisit the use of dwelling-level electricity data for the 2026 Census.

For transparency, IIS has, in this section, retained much of the initial discussion and findings regarding electricity usage data. This information may continue to have relevance for any future decisions to collect and use dwelling-level electricity data.

7.1 Electricity usage data and smart meters

At the time of IIS's initial PIA analysis, ABS was working with energy distributors to confirm what data it would collect. Variables were likely to include NMI (national meter identifier), address (or encoded address) connection status, time stamp, usage across specific period. Usage data would vary depending on whether the residence has a smart meter or a traditional meter. Currently, Victoria has close to full smart meter roll out across households in the state. The rest of the country is at about 10-20% smart meter penetration. The figure in NSW is set to rise with that state installing smart meters as traditional meters come up for replacement. (Meters have a 10-15-year lifespan.)

The granularity of the data is very different between regular and smart meters. IIS understands that regular meters record electricity usage on a quarterly basis whereas smart meters record usage in 30-minute increments. This is likely to move to 5-minute increments in the future. Fine grain data, of course, increases the privacy impact.

7.2 Is electricity usage data personal information?

Energy distributors appeared to believe that the usage data was not personal information. If it were, then disclosure to the ABS would be blocked by APP 6. However, some stakeholders thought that the NMI, which can be linked to address, would render the data 'personal information' as defined in the Privacy Act.

Determining whether data is personal information can be difficult and is affected by context. ABS made clear that it was not proposing to collect 'consumer' or 'retail' data (such as customer name or bill payment history). Address data will generally be personal information unless it has been encoded and separated from other identifiers appropriately. Whether the NMI (and usage data) is personal information is also context-dependent. It may be non-identifying alone (depending on the context) but become identifying when associated with other types of data.

If ABS decides to pursue use of electricity data in the future, it should work with energy distributors to ensure that the data they exchange is not personal information and does not become personal information during processing, otherwise the APPs will apply (and block data sharing). Where data is or can be linked to an address, (including via an NMI) it will meet the definition of personal information. This is the case even where there is more than one person living at the address. The OAIC points out that 'personal information of one individual may also be personal information of another individual...'²² While ABS may put measures in place to prevent linkages and to 'unidentify' electricity usage data, even small re-identification risks may have large consequences (including Privacy Act non-compliance). This will be something to manage with care for any future use of dwelling-level electricity data.

7.3 Privacy impact and stakeholder feedback

Although ABS may find ways to exchange electricity data in a form that does not identify individuals or their addresses, some major privacy concerns remain, and this was evident in stakeholder feedback. One person said they were 'horrified about collection of energy usage' and asked if ABS was informing people about this. They pointed out that 'when I signed up with my energy company, I did not sign up for this.'

Other stakeholders also expressed concern, saying: 'Smart meter data is about as sensitive as it gets,' and 'Collecting energy data sounds really shady.' Some thought that electricity data should not be used at all and that the benefits of the use of such data was disproportionate to the privacy impact and was likely to erode trust. One stakeholder commented that 'the conflation of consumer datasets with government datasets is very worrying.' However, this view was not shared by all with another stakeholder commenting: 'I personally see no issue with using non-government administrative data to increase the quality of the Census.'

Some of those who were concerned about the collection of electricity usage data felt that individuals should have the ability to opt-out of their data being collected and used in this way.

Stakeholders also wanted to know whether ABS knew how much better the data would be with or without the electricity usage data. One stakeholder asked whether ABS had a decision process to decide whether the data provides enough value or whether a particular dataset was worth using. Another said that ABS was externalising the privacy cost and this cost was borne by people at risk. They observed that if ABS counted the privacy cost, the economic verdict would be different. Other comments on this point included:

- 'The case hasn't been made about the level of the problem and that it merits the involvement of this kind of data. I'm not sure the trade-offs stack up with the problem.'
- "Could improve' does not justify the privacy invasion.'

²² OAIC, *APP Guidelines*, paragraph B.87.

KEY RISK AREA 3 – Electricity usage data

- 'It's not clear what effect use of electricity data has on the overall picture. Is it a big gain or a small gain? What's the effect on precision from a big picture perspective? What is the actual impact? If the gains are small, then it might not be proportional to the privacy impact.'
- 'ABS needs to document the level of improvement. Its analysis should be peer-reviewed and replicable. Having a strong evidence base is important before proceeding.'

Some stakeholders were less concerned about use of (non-government sourced) admin data. One data user pointed out: 'Most of our projects start with the Census data and we build on from there. For our work having accurate baseline data is essential. Extending the reach of data outside of government might erode a bit of trust in the ABS but reduced [data] quality will for sure reduce trust. Having some announcements made around what data will be used would suffice for me.'

Some stakeholders also speculated about how issues such as residential use of solar or apartment blocks wired to a single meter might reduce data utility too much. One stakeholder commented: 'There's a fudge factor here. This is about needing precise data. It seems like a lot of effort and privacy impact for not very accurate data.'

These comments demonstrate that, even if privacy risks are reduced by safeguards ABS puts in place, stakeholders are concerned about collection and use of electricity data by ABS. For future uses of dwelling-level electricity data, ABS will need to proceed cautiously and build trust. It should also make the case for why this data is necessary and worth the privacy impact before it proceeds.

Recommendation 10 – Conduct and publish further in-depth analysis of the level of improvement to Census data the use of electricity data would bring about before proceeding with any future uses

For any future proposals to use electricity usage data, ABS should conduct and publish in-depth analysis of the level of the problem, how much of a difference use of electricity data would make and whether the improvement in statistical accuracy is large enough to justify the privacy impact. Where the privacy impact is too high, ABS should not proceed or should explore options to lessen the privacy impact, such as the use of an opt-out mechanisms to allow individuals to opt-out of having their electricity usage data collected by ABS.

7.4 Data minimisation

ABS proposed collecting admin data (including electricity usage data) for the full population rather than only the addresses that did not respond to the Census. IIS discusses the issue of data minimisation in [section 8.2](#) below.

8. Other considerations

8.1 Sensitive information

Use case 1 will involve collection and use of 'Indigenous status' (that is, whether a person is Aboriginal or Torres Strait Islander). It will also collect information about individual's country of birth which may reveal information about a person's ethnicity. Both of these types of information are 'sensitive information' for the purposes of the Privacy Act and some extra protections apply.

It is worth noting that ABS has used this sort of information in the past for the same purpose – the difference being that previously it sourced the data from previous Census data, whereas this time it intends to source it from the SSRI dataset. The reason for this is that SSRI data will be more current than the 5-year-old Census data. For that reason, risks – such as any risks associated with identifying areas where subpopulations are concentrated – are not new for this project.

IIS finds that ABS's collection and use of this information meets the requirements of APP 3 and 6. And after the initial collection, the data will be aggregated to meshblock level, which lessens the privacy impact further. IIS does acknowledge concern from stakeholders, however, for the potential for use case 1 or associated data to be used for profiling to target subpopulations such as Aboriginal and Torres Strait Islander people. Some stakeholders also pointed out that certain refugee and migrant populations may be sceptical about surveillance and use of admin data about them. ABS should continue to engage with stakeholder groups representing these subpopulation groups. See [section 5.4](#) and recommendation 2.

8.2 Collection (APP 3)

IIS finds that collection of the admin data is authorised under ABS's enabling legislation (see [section 3.2](#) above).

Use case 1 requires data for the full population to meet its objectives. However, use cases 2 and 3 would target only those households that did not return a Census form. Therefore, there is a question about data minimisation and whether ABS is meeting the 'reasonably necessary' requirement contained in APP 3.²³ ABS has indicated that it will not proceed with the use of dwelling-level electricity usage data in Census, so this point may be moot. However, if it decides to go ahead with using such data in the future, it should consider ways to minimise the amount of data to be collected. APP 3 may not strictly apply to electricity usage data if it does not meet the definition of personal information. That said, stakeholders were particularly concerned about ABS's use of this data so complying with APP 3 may offer additional reassurance to the community that privacy best practice is being applied.

²³ APP 3 states that an agency may only collect personal information that is reasonably necessary for, or directly related to, one or more of its functions or activities, see APP 3.1.

Prior to ABS's decision not to go ahead with use of dwelling-level electricity data, stakeholders discussed data minimisation and their discussions revealed that the matter was a complex one. With regard to electricity data in particular, a number of stakeholders said that collecting data for all Australian households (rather than a subset) constituted a significant overreach on the part of ABS and clashed with data minimisation principles.²⁴ Stakeholders wanted to know that ABS would only collect the data necessary – in this case, data relating to households that did not return a Census form. Data custodians pointed out that this created a new privacy issue, in that ABS would have to disclose to data custodians the list of non-responding addresses in order to receive back a smaller dataset. Moreover, having data custodians conduct some of this data analysis first shifted a bigger data processing burden onto the custodian. Others thought that there may be more trust in the community for ABS to conduct the data processing from the full dataset. This feedback arose particularly in relation to electricity data, but the same considerations arise for other admin datasets.

Assessing what is appropriate from a data minimisation standpoint is clearly not straightforward. Both approaches (collecting all or collecting only some of the data from data custodians) have potential privacy impacts. Initially, IIS considered recommending that ABS only collect a reduced dataset where possible. However, given the complexity of the matter and the secondary privacy issues that arise, IIS believes this is a matter that requires ABS's data processing expertise to determine whether alternative strategies exist to enable collection of a reduced dataset in a way that offsets secondary privacy issues. Additionally, with ABS deciding not to pursue use of dwelling-level electricity data, this consideration is less immediate. That said, it is something to return to if ABS pursues use case 3 or if it decides to use dwelling-level electricity data in conjunction with future Censuses.

Recommendation 11 – Explore options to reduce the amount of data collected in the event that ABS pursues use case 3

ABS should determine whether alternative strategies exist to enable collection of a reduced dataset (that is, only data associated with addresses that did not respond) in a way that offsets secondary privacy issues (such as issues arising from having to disclose addresses that did not respond). ABS should take into account the practicability of alternative options (including whether they would impose an unreasonable processing burden on data custodians).

8.3 Use (APP 6)

Regarding use case 1, IIS finds that there are rigorous processes in place to ensure data is used appropriately during integration. Once data is aggregated to the meshblock level (which is the intention for use case 1), APP 6 is unlikely to apply (so long as the data no longer meets the definition of personal information). For use cases 2 and 3, IIS finds that ABS has rigorous processes in place to ensure data is subject to strict governance controls during data integration and use. Each step is

²⁴ There was also concern among stakeholders about ABS collecting energy usage data over a period of time on either side of Census night, rather just Census night itself. One stakeholder asked: 'why is it necessary to get a detailed record of energy? Wouldn't it be better to get the minimal amount of data – for example the distributors tell you who looks like they were home.'

clearly defined along with the roles of staff. Moreover, ABS's use of the data for this project aligns with its authorities under its enabling legislation.

A risk to manage is that of scope creep. Some of the data that ABS may choose to use in the future, such as dwelling-level electricity data, offer a rich new source of insight into possible novel areas of inquiry. Having strict governance along with clear data disposal schedules will help avoid unauthorised or inappropriate additional uses. Needless to say, if ABS decides to use any of the admin datasets (assessed in this PIA) for other uses, it should undertake another PIA.

8.4 Disclosure (APP 6)

ABS will be using integrated admin data internally. It does not have plans to disclose it to external recipients. Therefore APP 6 disclosure provisions largely do not apply.

There were some questions from stakeholders about whether data custodians would get consent to disclose the data to ABS (or, alternatively whether ABS would get consent to collect the data). Consent is not required in this scenario as data custodians are able to disclose the data to ABS under the legislative authorities outlined in [section 3.3](#). The data custodians that must take additional care in their disclosure of data to ABS are energy distributors. If the data is personal information, the Privacy Act (and consent obligations) will apply to energy distributors that are also APP entities. State-based public sector bodies may be covered (in a similar way) by state-based privacy law. (APP 6 is unlikely to apply to electricity data aggregated to the meshblock level if it no longer meets the definition of personal information.)

While consent is not a requirement under the Privacy Act for the Census admin data project, adopting an opt-out mechanism could be an option for aspects of the data sharing where the privacy impact is, or may appear to the community to be, high. See recommendation 10.

8.5 Security (APP 11)

APP 11 requires an agency to take reasonable steps to secure the personal information it holds. It also requires an agency to dispose of personal information once the information is no longer needed for any purpose for which the personal information may be used or disclosed under the APPs.

ABS has a range of strong security arrangements in place for the data it handles. Data integration (the librarian, linker and assembler steps outlined in [Appendix C](#)) occurs in the ABS's Next Generation Infrastructure which, according to ABS, provides secure, restricted access. Role-based access controls are implemented according to functional separation principles. It also provides audit trails of access to the data. Analysis and statistical use of the data occurs in the Census data processing environment. IIS understands this to be a secure Oracle-based environment with role-based access controls.

The security settings for ABS's data integration activities have been assessed in detail in the earlier MADIP PIAs and will not be changing for this project. IIS has not conducted a security audit of ABS's arrangements applying to this project, but became familiar with security arrangements from material provided by ABS including the MADIP PIAs. IIS encourages ABS to continue to take a rigorous approach to security, including regular auditing. We did not identify any further issues.

Regarding data disposal requirements, we note that the draft data integration plan specifies that data will be retained until there is no longer a statistical need for the purposes outlined in the plan, or until the end of 2022, whichever is earliest. It is helpful to have a clear statement to this effect recorded in the plan. ABS should also clarify if raw data has the same retention period. IIS suggests also making retention information clear in public-facing material, such as the privacy policy (see recommendation 5).

9. Appendix A – Glossary

Term or acronym	Meaning
ABS	Australian Bureau of Statistics
APPs	Australian Privacy Principles (a set of 13 rules in the Privacy Act)
ARID	Address Register ID (identifier used to represent addresses, often further encoded to reduce identifiability of datasets)
De-identified data	<p>Personal information is de-identified 'if the information is no longer about an identifiable individual or an individual who is reasonably identifiable' (section 6(1) of the Privacy Act).</p> <p>See also the meaning of unidentified data.</p>
Disposal schedules	<p>Timeframes for disposing of (or permanently deleting) data.</p> <p>Under the Privacy Act, entities must dispose of personal information once the information is no longer needed for any purpose for which the personal information may be used or disclosed under the APPs. Generally, entities meet this obligation by implementing data retention policies which establish schedules for data disposal.</p>
Five safes framework	<p>This framework takes a multi-dimensional approach to managing disclosure risk. The five elements of the framework are: safe people; safe projects; safe settings; safe data; and safe outputs. The framework poses specific questions to help assess and describe each risk aspect (or safe) in a qualitative way. Controls can be placed on the data and the manner in which data are accessed. The framework is designed to facilitate safe data release and prevent over-regulation.</p>
IIS	Information Integrity Solutions Pty Ltd
MADIP	Multi Agency Data Integration Partnership
MCD	Medicare Enrolments Database (a MADIP dataset to be used in the Census admin data project)
Meshblock	The smallest geographic area in the Australian Statistical Geography Standard. Most meshblocks contain 30 to 60 dwellings.
NDAC	National Data Acquisition Centre (collection point for datasets from external data custodians)
NGI	Next Generation Infrastructure (the IT environment ABS uses for data integration and processing)
NMI	National Meter Identifier
OAIC	Office of the Australian Information Commissioner
ONDC	Office of the National Data Commissioner

Term or acronym	Meaning
PIT	Personal Income Tax Client Register (a MADIP dataset to be used in the Census admin data project)
Privacy Act	<i>Privacy Act 1988</i> (Cth)
SSRI	Social Security and Related Information (a MADIP dataset to be used in the Census admin data project)
Unidentified data	Data is considered 'unidentified' when direct identifiers such as name and address are removed or altered into an unidentifiable form. Further confidentialisation or safeguards are often required for the data to be considered de-identified. See also the meaning of de-identified data.

10. Appendix B – Data variables involved

As explained in the report, ABS separates ‘linkage variables’ from ‘analytical variables’. For use cases 1 and 2, the linkage variable is the address which is converted into an ‘address register ID’ or ARID – a code that stands in for the address. The ARID is then further encoded.

As agreements with data custodians are formalised, there may be changes to the variables involved or how they are named. This table is indicative only.

Dataset	Analytical variable	
SSRI	Address start date	
	Address end date	
	Address type	
	Geography of address (Meshblock, SA1, SA2, SA3, SA4, Greater Capital City Statistical Areas, State/territory, Remoteness Areas)	
	Age	
	Sex	
	Benefit status	
	Benefit type	
	Country of birth	
	Date of death	
	Indigenous code	
	MCD	Address start date
		Address end date
Address type		
Geography of address (Meshblock, SA1, SA2, SA3, SA4, Greater Capital City Statistical Areas, State/territory, Remoteness Areas)		
Entitlement type		
Entitlement start date		
Entitlement end date		
Entitlement country code		
Consumer start date		
Consumer end date		
Medicare card expiry date		
Status of Medicare card		
Departure date		
Age		
Sex		
Date of death		

Dataset	Analytical variable
PIT	Address start date
	Address end date
	Address type
	Geography of address (Meshblock, SA1, SA2, SA3, SA4, Greater Capital City Statistical Areas, State/territory, Remoteness Areas)
	Age
	Sex
	Deceased estate
	Number of dependent children
Electricity usage data	Meter number (NMI)
	Connection status
	Time stamp
	Quality of read
	Total usage across specific period

11. Appendix C – Information flows and data separation

11.1 Collection

The ABS receives admin datasets via its National Data Acquisition Centre (NDAC). In cases where the dataset is too large to receive via the NDAC, it is brought into ABS on an encrypted portable storage device. Measures are in place to ensure both the security of the device and that the device is wiped once the data is uploaded to ABS's secure environment. As noted above, the data arrives in ABS's environment with linkage variables separated from analytical variables. Then the data is moved to the data integration Next Generation Infrastructure.

11.2 Use

In NDAC environment, the data is anonymised, linked, assembled and analysed, with data separation principles deployed to ensure that no single ABS officer has a full view of the data. The purpose of data integration is to link two or more separate datasets together. Data processing is conducted by four separate people:

Librarian

The librarian receives just the datasets containing the linkage variables (in this case the addresses associated with the analytical dataset) and not the analytical variables. They clean the data and code the addresses to address register IDs (ARID) and further anonymise the data through hashing. Then they pass the anonymised datasets to the assembler. The librarian is the only person in a project who sees the linkage variable (ie, the addresses). However, they never see the connections with analytical datasets, so while they do get access to personal information, they get no insight into linkage or analysis.

Linker

Sometimes a linker is involved. They receive the anonymised datasets of linkage variables and run them through a linking process to find matches and then produce a concordance report – basically a report that states that record X in dataset A links to record Y in dataset B. Generally, the linker only handles the linkage variables. For this project, IIS understands that a linker will not be needed as all the datasets will be merged using the ARID.

Assembler

The assembler is the first person who receives and handles the analytical variables. They take the analytical datasets and the hashed ARID datasets and assemble (or merge) the data. For this project the assembler then:

- Creates an extract for enumeration and planning aggregated to the meshblock level and passes this to the Census Enumeration team
- Creates an extract for occupancy, runs a probability model and passes the resulting output aggregated to hashed ARID to the Census Processing team
- Creates an extract for donor selection aggregated to hashed ARID and passes it to the Census Processing team.

Analyst

The analysts use the data. IIS understands that analysis work will be undertaken by staff in the Census Futures, Household Statistics Methodology, Methodology Futures, Census Enumeration and Census Data Operations sections.

11.3 Disclosure

The integrated data is not disclosed outside of the ABS. This is the central difference between this project and other projects under MADIP where the purpose of the integration is to make the data available externally to researchers.

11.4 Disposal

The data integration plan specifies retention arrangements for particular projects. For use cases 1 and 2, IIS understands that the integrated data will be securely disposed of by the end of 2022 or when no longer needed to achieve project objectives – whichever comes first.

Raw admin data – the original dataset before data integration – is retained or disposed of depending on arrangements agreed with the data custodian. Certain datasets are classed as 'enduring' including those used to create the Person Linkage Spine, as ABS reviews these datasets each year. Datasets may also be classed as non-enduring in which case they are deleted (with the processed data) after the project is complete.

12. Appendix D – APP compliance table

This table gives a summary of APP compliance considerations for the Census admin data project. Note that not all recommendations are represented in this table. This includes, for example, some recommendations that go to managing social licencing issues and fostering trust which do not line up with a particular APP.

APP	Compliance considerations	Best practice considerations
1 Openness	<p>All use cases</p> <p>Update privacy policy. See recommendation 5.</p> <p>IIS finds ABS has strong governance arrangements in place to help it to meet the requirements of the APPs.</p> <p>No further issues identified.</p>	<p>All use cases</p> <p>Include information about retention of integrated admin data in the privacy policy. See recommendation 5</p> <p>Enhance measures to monitor accumulating impact on privacy. See recommendation 1.</p> <p>Develop and deploy a communications plan for the Census admin data project and engage civil society. See recommendation 7.</p> <p>Use case 2</p> <p>Publish research that explains the return in terms of improvement of Census data gained from use of electricity usage data. See recommendation 10.</p>
2 Anonymity	<p>All use cases</p> <p>ABS will not be engaging directly with data subjects for this project; therefore, APP 2 is not relevant here. IIS also notes that ABS will not be collecting name information and will largely use data with identifiers removed and linkage variables encoded.</p> <p>No issues identified.</p>	

APP	Compliance considerations	Best practice considerations
3 Collection of solicited PI	<p>All use cases</p> <p>IIS finds that collection of the admin data is authorised under ABS’s enabling legislation.</p> <p>Collection infrastructure (including the NDAC) and other acquisition governance appears to be strong.</p> <p>IIS notes that use case 1 requires data for the full population to meet its objectives. However, use cases 2 and 3 intend to target only those households that did not return a form. ABS should check its collection of personal information reasonably necessary to achieve its purposes for use case 2. See recommendation 11.</p>	<p>Use case 2</p> <p>Conduct research that quantifies the level of improvement of Census data gained from use of electricity usage data. Assess whether the return is great enough to justify the privacy impact. Conduct this assessment before proceeding with collection. See recommendation 10.</p>
4 Collection of unsolicited PI	<p>All use cases</p> <p>ABS will not be collecting unsolicited personal information for this project. Data variables are decided with data custodians in advance.</p> <p>No issues identified.</p>	<p>All use cases</p> <p>IIS agrees with suggestions made in the MADIP PIA update that ABS should work with data custodians to minimise the risk that unsolicited data is included in datasets provided to ABS.</p>
5 Notice	<p>All use cases</p> <p>ABS should continue to work with data custodians to manage compliance with APP 5. See recommendation 6.</p>	

APP	Compliance considerations	Best practice considerations
<p>6 Use</p>	<p>Use case 1</p> <p>IIS finds that ABS has rigorous processes in place to ensure data is subject to strict governance controls during integration. IIS notes that use case 1 involves aggregation of integrated data to the meshblock level. We find that, at that point, APP 6 is unlikely to apply, where the data no longer meets the definition of personal information.</p> <p>Use cases 2 and 3</p> <p>IIS finds that ABS has rigorous processes in place to ensure data is subject to strict governance controls during data integration and use. Each step is clearly defined along with the roles of staff. Moreover, ABS’s use of the data for this project aligns with its authorities under its enabling legislation.</p> <p>A risk to manage is that of scope creep. Having strict governance along with clear data disposal arrangements will help avoid unauthorised or inappropriate additional uses. Needless to say, if ABS decides to use any of the admin datasets (assessed in this PIA) for other uses, it should undertake another PIA.</p> <p>No further issues identified.</p>	<p>Use case 3</p> <p>Before using admin data to repair the Census, establish threshold tests that allow ABS to determine when damage to Census data is significant enough to warrant pursuing use case 3. See recommendations 3 and 4.</p>
<p>6 Disclosure</p>	<p>IIS understands that ABS does not intend to disclose admin data outside of ABS. The three use cases involve internal use only. For that reason, IIS has not identified any issues related to APP 6 for ABS.</p> <p>Data custodians disclosing data to ABS do so under legislative authorities outlined in section 3.3 of the report.</p>	
<p>7 Direct marketing</p>	<p>All use cases</p> <p>APP 7 applies to private sector organisations rather than public sector agencies.</p> <p>No issues identified.</p>	

APP	Compliance considerations	Best practice considerations
8 Cross border disclosure	<p>All use cases</p> <p>Integrated admin data will be used internally and will not be disclosed outside of ABS.</p> <p>No issues identified.</p>	
9 Identifiers	<p>All use cases</p> <p>ABS will not be collecting and using government issued identifiers for this project.</p> <p>If ABS opts to collect and use government identifiers for use case 3 it should ensure it does so in compliance with APP 9.</p> <p>No issues identified.</p>	
10 Data quality	<p>All use cases</p> <p>IIS finds that the data processing for the Census admin data project is aimed at improving the accuracy of Census data. Therefore, it can be considered to have an overall positive impact on data quality.</p> <p>ABS cannot control the level of quality of the data sets that are provided to it. However, through linkage across datasets, ABS is able to refine the accuracy of the data.</p> <p>It should also be noted that, following processing, the data is rendered, if not de-identified then, un-identified and is not used in a way that will have a direct one-to-one impact on data subjects.</p> <p>No issues identified.</p>	
11 Security	<p>All use cases</p> <p>Continue to apply security arrangements in place for ABS data integration activities. Continue with regular security audits.</p> <p>Clarify retention arrangements for integrated admin data in public-facing material. See recommendation 5.</p>	<p>All use cases</p> <p>IIS agrees with suggestions in the MADIP PIA update that ABS should be transparent about security arrangements by offering details online.</p>

APP	Compliance considerations	Best practice considerations
12 Access	<p>All use cases</p> <p>ABS has a general exemption to access requests relating to personal information it collects under the Census and Statistics Act. This is because APP 12 allows an agency to refuse access where they are authorised under Freedom of Information legislation to do so. The Freedom of Information Act enables refusal on this ground.</p> <p>No issues identified.</p>	
13 Correction	<p>Use case 3</p> <p>Where ABS decides to collect and impute data at the person level and that data meets the definition of personal information, ABS will need to ensure its internal correction procedures enable it to act on correction requests. This may be a matter to explore in the use case 3 PIA.</p>	

13. Appendix E – PIA methodology

IIS conducted the PIA in a series of phases, including:

- **Planning**

In this phase IIS confirmed the scope and objectives of the PIA and confirmed the tasks, milestones and timeframes for the project. IIS also worked with ABS to plan consultation with stakeholders.
- **Information gathering**

In this phase, IIS gathered information to ensure a sufficient understanding of the ABS's planned use of administrative data – including the nature of the personal information involved, and how it would be stored, managed and used. IIS gathered information both through reading documentation and meeting with ABS staff. Documents that IIS reviewed are listed at [Appendix F](#).
- **Consultation with external stakeholders**

IIS attended three roundtable sessions with external stakeholders. The sessions were arranged and run by ABS and included stakeholders representing data custodians, researchers, civil society and privacy advocacy. During these sessions, IIS listened to feedback about the Census admin data project, areas of concern and options for reducing privacy impact. Stakeholders consulted for this PIA are listed at [Appendix G](#). Further information about the consultation is provided in [section 2.2](#) of the report.
- **Analysis**

In this phase, IIS identified relevant privacy issues and options to mitigate risks identified. This included assessing project information flows against the APPs and considering broader issues around best practice, building trust and managing risks of scope creep.
- **Drafting the PIA report**

In this phase, IIS wrote up its analysis and recommendations into the report you are reading. We also produced a short summary report of stakeholder feedback arising in the three roundtable sessions. IIS presented the draft PIA report to ABS for review and feedback and then finalised the document, taking that feedback into account.

14. Appendix F – Documents reviewed

Document title / name

PIA – Linkage of address register with Census data Privacy Impact Assessment 2017

PIA – MADIP Independent Privacy Impact Assessment 2018

PIA – MADIP Independent Privacy Impact Assessment 2018 – response by MADIP agencies

PIA – National Health Survey Linkage Project Independent Privacy Impact Assessment 2018

PIA – National Health Survey Linkage Project Independent Privacy Impact Assessment 2018 – response by ABS

PIA – MADIP Privacy Impact Assessment Update 2019

PIA – MADIP Privacy Impact Assessment Update 2019 – Consultation report

PIA – MADIP Privacy Impact Assessment Update 2019 – Maddocks Independent Assurance Report

PIA – MADIP Privacy Impact Assessment Update 2019 – response by MADIP Board

PIA – Stats NZ Creating the 2018 Census dataset by combining administrative data and census forms data: Our privacy impact assessment 2019 (Second edition)

ABS webpage – [Planning the 2021 Census](#)

ABS webpage – [Census of Population and Housing: Census Dictionary, 2016: Derivations and imputations](#)

ABS webpage – [Administrative Data Research for the 2021 Census](#)

ABS webpage – [Can administrative data help to improve the Census count?](#)

Independent Assurance Panel: Report on the quality of 2016 Census data, 2017

Australian Statistics Advisory Council – meeting paper and PowerPoint presentation – *Community Trust and Data Ethics* – 13 November 2019

Building trust in the ABS and our data use strategy

Background information on use of integrated administrative data for the 2021 Census

Background information on use case 3

ABS diagram – *Basic principle of functional separation*

ABS diagram – *Data flow – Census dwelling asset*

ABS diagram – *Key data integration governance*

Draft data integration plan – Using administrative data to improve and support the 2021 Census of Population and Housing

Document title / name

Article – 'NZ census gaps to be filled with 'admin data' and will be reliable, says Stats NZ', Stuff.co.nz, 22 August 2019.

ABS Data Linkage Centre Incident Response Plan

MADIP PIA and Census PowerPoint presentation

Valuing the Australian census – Lateral Economics – 27 August 2019

15. Appendix G – Stakeholder meetings held

Date	Participating stakeholder groups
Roundtable 1 – 10 February 2020	<ul style="list-style-type: none"> ● Australian Institute of Health and Welfare ● Australian Taxation Office ● C4NET (Electricity provider) ● Department of Environment, Land, Water and Planning (Victoria State Government) ● Department of Health ● Department of Industry, Innovation and Science ● Department of Social Services ● Office of the National Data Commissioner, Department of the Prime Minister and Cabinet ● Services Australia
Roundtable 2 – 14 February 2020	<ul style="list-style-type: none"> ● Centre for Big Data Research in Health, University of New South Wales ● Health Services Research Association of Australia and New Zealand ● Office of the Australian Information Commissioner ● Queensland Council of Civil Liberties and Electronic Frontiers Australia
Roundtable 3 – 17 February 2020	<ul style="list-style-type: none"> ● Access Now ● Deakin University ● Digital Rights Watch ● Griffith University ● Liberty Victoria ● Melbourne Institute of Applied Economic and Social Research, University of Melbourne ● NSW Council of Civil Liberties ● People with Disability Australia ● Qld Office of the Information Commissioner ● School of Computing and Information Systems, University of Melbourne



**INFORMATION
INTEGRITY
SOLUTIONS**

Information Integrity Solutions Pty Ltd

PO Box 978, Strawberry Hills NSW 2012, Australia

P: +61 2 8303 2438

F: +61 2 9319 5754

E: inquiries@iispartners.com

www.iispartners.com

ABN 78 107 611 898

ACN107 611 898