

Privacy Impact Assessment on the Linkage of Address Register with Census Data/Contact Data

For: Australian Bureau of Statistics

Date: Final – Updated 31 October 2017

INFORMATION
INTEGRITY
SOLUTIONS

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

Table of Contents

1. Introduction	3
2. Background	4
2.1 Address Register	4
2.2 Proposal 1 - Linking Addresses with Census Data (for More Focused Sampling)	6
2.3 Proposal 2 - Linking Addresses with Contact Data (for Calling Respondents)	7
2.4 Contact Data	8
3. Methodology	9
3.1 Context for the PIA	9
3.2 Process	9
4. Proposed Information Flows	10
5. Positive Privacy Impacts	11
6. Privacy Risks	12
6.1 Social Licence and Public Perception	12
6.2 Transparency Issues	13
6.3 Opportunity to Consent	14
6.4 Consultations on the Proposed Linkages	15
6.5 Engagement and Consultation with Population Groups of Increased Interest	15
6.6 Data Security, Retention and Access	16
7. Recommendations	18
7.1 Recommendation 1 – Social Licence and Transparency	18
7.2 Recommendation 2 – Consultations with Communities of Increased Interest	19
7.3 Recommendation 3 – Pilot Testing	19
7.4 Recommendation 4 – Unequal Sampling Policy	20
7.5 Recommendation 5 – Address Register and Address Register IDs	20
7.6 Recommendation 6 – Opportunity to Opt-Out of Future Phone Contact	21
7.7 Recommendation 7 – Compliance with APPs	21
1. Appendix 1 - Documents Reviewed	22
2. Appendix 2 - Interviews Held (9 Aug 2017)	24
3. Appendix 3 - Compliance Risk Table	25

1. Introduction

The Australian Bureau of Statistics ('ABS') purpose is to inform Australia's important decisions by partnering and innovating to deliver relevant, trusted, objective data, statistics and insights. As the central statistical authority for the Australian Government and provider of statistical services to the states and territories, the ABS seeks to deliver the most public value it can from available resources. A key function of the ABS is to provide the statistical information that Australia needs. For some populations, such as Aboriginal and Torres Strait Islanders, increasing the quality and availability of trusted official statistics is important to inform public debate and enable decision-making.

The ABS proposes to link the ABS Address Register ('Address Register') with data obtained from the Census of Population and Housing ('Census Data') (Proposal 1) and data related to telephone numbers obtained from external sources ('Contact Data') (Proposal 2). These proposals align with the broader ABS Transformation program which includes modernising ABS infrastructure and business processes to reduce risks to statistical outputs; reduce costs; reduce red tape for providers; and achieve faster turnaround in dissemination of statistics. A key focus of the transformation is to enhance ABS' statistical capability by redesigning the statistical collections, methods, processes, products and services in order to deliver a more sustainable statistical program.

The purpose of the address-census data linkage proposal is to allow ABS to locate populations of interest more efficiently to:

1. enhance the quality and frequency of statistics produced for decision making particularly in relation to key population groups of interest;
2. enable the ABS to approach fewer dwellings to find respondents of interest (in scope) when conducting the ABS Household Survey Program, reducing the impact on people who would otherwise be contacted and
3. may reduce cost for the acquisition of data.

The proposed addresses-contact numbers linkage would enable ABS interviewers to call potential respondents by telephone to improve the efficiency and effectiveness of the ABS Household Survey Program.

The proposed linkages are still in an early investigative stage. A large part of the proposed linkages' success will depend on whether they meet legislative privacy requirements and community privacy expectations. As such, the ABS has engaged Information Integrity Solutions Pty Ltd ('IIS'), a privacy consultancy, to conduct a Privacy Impact Assessment ('PIA') of the two proposals.

This PIA:

- maps the data flows with respect to the two proposals
- examines the privacy considerations around the two proposals
- evaluates whether the two proposals would meet community privacy expectations and obligations under the *Privacy Act 1988* (Cth) ('Privacy Act')

- makes recommendations on appropriate mitigation strategies with respect to the two proposals where key privacy risks are identified.

2. Background

Subject to the *Australian Bureau of Statistics Act 1975* (Cth) and the *Census and Statistics Act 1905* (Cth), the ABS has the power to collect, compile, analyse and disseminate statistics and related information;¹ it may also from time to time collect such statistical information in relation to the matters prescribed in legislation it considers appropriate.²

Other than the Census of Population and Housing ('Census') held every five years, the ABS has been gathering information from individuals and businesses through different surveys, including the Monthly Population Survey, the National Health Survey, the Survey of Income and Housing and the National Aboriginal and Torres Strait Islander Health Survey ('NATSIHS'). The statistics produced by these surveys are important across both business and government and cover the economy, society, environment and the population sectors and are vital in ensuring effective policy decisions and efficient use of government funds to achieve societal goals.

2.1 Address Register

The Address Register was established in 2015 for providing the ABS with a comprehensive list of all physical addresses in Australia. The Address Register has been developed for improving the efficiency and effectiveness of the Census of Population and Housing and household surveys, the linkages for ABS and external data linkage projects and the ability to conduct geospatial analysis.

The basis of the Address Register is the quarterly Geo-coded National Address File ('G-NAF') provided by PSMA Australia, which is supplemented with information from other available address data sources (such as Google Earth and Street view) and field work by ABS Officers, including the ABS canvassing officers. The Address Register consists of the following data items: Address Register ID, addresses, address 'Use' (whether it is commercial or residential), Geocode (latitude/longitude), dwelling structure, if an address is a Special Dwelling the type of special dwelling (e.g. hotel, hospital, etc), dwelling capacity (e.g. number of rooms or beds), or if located in an Aboriginal and Torres Strait Islander Community, the community name, number of dwellings, population count, whether it contains special dwellings and the governing organisation.

Access to the Address Register is currently provided to PSMA Australia for improving the quality of GNAF and GeoScience Australia for emergency response management.

¹ Section 6(b), Australian Bureau of Statistics Act 1975 (Cth)

² Section 9(a), Census and Statistics Act 1905 (Cth)

One question surrounding the Address Register is whether it contains personal information. The relevant definition as set out in the Privacy Act³ is:

‘information or an opinion about an identified individual, or an individual who is reasonably identifiable:

(a) whether the information or opinion is true or not; and

(b) whether the information or opinion is recorded in a material form or not.’

A person does not need to be named, or clearly identified, for information to be personal information. Information is ‘about’ an individual where there is a connection between the information and the individual,⁴ such as when identity may be established from the information by inference or knowledge.⁵

For example, in the NSW case, *Office of Finance and Services v APV and APW*,⁶ addresses in the conservation management plan and schedule of works (which also included information about the works carried out at the addresses, including photographs of the interior and exterior, floor plans and other internal design features) without containing the names nor any photographs of the individuals or other information were considered personal information. The Appeal Panel found that the critical factor was that it was possible to conduct a search on the internet of the address of their property and discover from the website the name of the individual. The Appeal Panel’s reasoning was as follows (at [56]):

“The definition of personal information states that the information is about an individual whose identity is apparent or can reasonably be ascertained from the information or opinion.” Those words do not mean that other material cannot be consulted. That is obvious from the fact that there are two ways in which information or an opinion may disclose a person’s identity. Either the identity is “apparent” from the information or it “can reasonably be ascertained” from that information. The dictionary definition of the adjective “apparent” is “capable of being clearly perceived or understood; plain or clear.” (Macquarie Dictionary online). The verb “ascertain” means “to find out by trial, examination, or experiment, so as to know as certain; determine.” (Macquarie Dictionary online). By including the option that a person’s identity can “reasonably be ascertained” from the information, the legislature was intending to allow a person to find out or determine the identity of the person from the information and, where reasonably identifiable from other information, from that other information.”

³ Section 6, Privacy Act 1988 (Cth)

⁴ See *Telstra Corporation Limited and Privacy Commissioner* [2015] AATA 991 (18 December 2015) at [112], and *Privacy Commissioner v Telstra Corporation Limited* [2017] FCAFC 4 (19 January 2017), at [43] and [64] per Kenny and Edelman JJ.

⁵ OAIC website, ‘What is personal information? - When is an individual ‘reasonably identifiable’?’, available at <https://www.oaic.gov.au/agencies-and-organisations/guides/what-is-personal-information>

⁶ See *Office of Finance and Services v APV and APW* [2014] NSWCATAP 88, available at <https://www.caselaw.nsw.gov.au/decision/54a63ffc3004de94513dc922>

Likewise, if information in the Address Register can be considered with other datasets that could ultimately or reasonably identify the individual, then that information in the Address Register would be considered personal information, as demonstrated in the above case. That same reasoning would apply to the definition of personal information in the Federal Privacy Act. The NSW case is particularly useful as it provides a similar address scenario. Determining whether a person is 'reasonably' identifiable will require a contextual consideration of the particular circumstances, depending on the cost, difficulty and likelihood that the information will be connected in such a way.

While this PIA does not try to provide any legal advice on this question, it points out that in some contexts the Address Register will constitute personal information. For instance, where the ABS has a large number of people being able to access it, and other data points are associated with it, it is more likely to constitute personal information.

Furthermore, according to the Office of the Australian Information Commissioner (OAIC), where there is uncertainty, entities are encouraged to err on the side of caution by treating the information as personal information, and handle it in accordance with the Australian Privacy Principles (APPs).

It is also worth mentioning that according to the Australian Community Attitudes to Privacy Survey 2017, addresses are considered as one of the top three pieces of information Australians are most reluctant to provide.⁷

2.2 Proposal 1 - Linking Addresses with Census Data (for More Focused Sampling)

Census Data is collected compulsorily from all individuals in Australia during the Census Period every five years. Census Data details a number of personal and household attributes, which is currently used for the compilation of statistical data for use by the government, organisations and individuals.

The first proposal is to link certain characteristics of Census Data collected from the 2016 Census with the addresses in the Address Register. Such linkage can reveal characteristics of particular dwellings (at the time of the Census), including whether a household contained at least one usual resident of Aboriginal or Torres Strait Islander origin, or whether a household contained at least one usual resident of a certain age, sex, income ranges, migration status and disability, etc., at the time of the Census. With such information, the ABS would be able to change its current methodology of choosing samples of dwellings in areas using Random Sampling (i.e. choosing dwellings randomly) to a more focused sample selection that increases the probability of selection (i.e. Unequal Probability of selection) of certain dwellings within areas by varying the chance of choosing dwellings that were more likely to contain a person with the characteristics of a particular population of interest based on information at the time of census.

⁷ OAIC, Australian Community Attitudes to Privacy Survey, available at <https://www.oaic.gov.au/engage-with-us/community-attitudes/australian-community-attitudes-to-privacy-survey-2017> p.i

This would significantly increase the likelihood of selecting dwellings that house people in population groups of interest to a particular survey, potentially saving ABS significant cost and reducing overall burden on the community.

For instance, in the 2012-13 National Aboriginal and Torres Strait Islander Health Survey ('NATSIHS'), the ABS had to approach and screen around 300,000 households in order to identify a sufficient number of dwellings housing an Aboriginal or Torres Strait Islander person to produce reliable statistics. This method places a burden on a large number of Australians and is time and resource consuming, as only 3% of the overall population are Aboriginal or Torres Strait Islander. For instance, in the 2012 NATSIHS, only around 9,000 out of the 300,000 households that were approached reported an occupant as being of Aboriginal or Torres Strait Islander origin.

Likewise, in the 2015-16 Household Expenditure Survey, 19,600 additional households were approached and screened to identify just over 2,400 households whose main source of income was government pension/benefits and/or allowances, to increase the reliability of the living cost index expenditure class estimates produced from the survey. The ABS thus proposes that a more focused selection of dwellings (i.e. unequal probability) within areas can be employed to improve the probability of selecting dwellings likely to house persons with characteristics of interest. The ABS is considering employing this method initially to the selection of dwellings likely to house Aboriginal or Torres Strait Islander persons for the upcoming NATSIHS, thus reducing the number of households the ABS needs to approach in total. For NATSIHS, the ABS is only considering using this approach in areas sparsely populated by Aboriginal and Torres Strait Islander persons; dwellings located in Discrete Aboriginal and Torres Strait Islander Communities or areas with a high concentration of Aboriginal and Torres Strait Islander persons will still be randomly selected.

The ABS also intends to consider the future integration of the Address Register with other dwelling level Census information to increase the efficiency of other surveys such as the Household Expenditure Survey. However, this PIA will focus particularly on NATSIHS as that is the potential first implementation of this revised approach to surveying. It is expected that the recommendations made in this report would be broadly applicable to other surveys. However, it will be important to consider any changing contexts and sensitivities that may arise in different surveys that may affect level of privacy risk.

2.3 Proposal 2 - Linking Addresses with Contact Data (for Calling Respondents)

The second proposed linkage concerns the linkage of addresses (in the Address Register) with Contact Data. Currently, when conducting Household Surveys, the ABS sends 'approach letters' to the selected dwellings asking them to complete a Household Contact Details Form online for scheduling interviews. However, the contact details form is completed by only 10-20% of selected dwellings. Therefore, in many instances, the interviewers approach dwellings in person without an appointment.

Access to the Integrated Public Number Database (IPND) (and other commercial sources such as Sensis) opens up the possibility of making contact by telephone which will significantly improve the

efficiency and effectiveness of the ABS Household Survey program. For instance, this could enable selected dwellings to be contacted by telephone to arrange an appropriate time to conduct the survey.

Additionally, the IPND/Sensis would be used to update address use on the ABS Address Register. This would be done by feeding into an address use predictive model, such as usage codes (business/ government/ residential/ charity etc), and assisting in identifying residential dwelling structure at residential addresses to determine whether an address is in scope of household surveys.

2.4 Contact Data

Contact Data here refers to contact numbers and other information (such as name) obtained from IPND and commercial sources, such as the Sensis White Pages.

The IPND is an industry-wide database containing all listed and unlisted public telephone numbers, and other information in relation to each public number of a customer, including the name of the customer, the service address of the customer, an alternate address, the type of service (e.g. fixed, mobile, payphone or premium services), etc.⁸

ABS has recently been authorised by the Australian Communications and Media Authority as a research entity to use IPND data to improve the quality of the Address Register.

At this stage, the ABS wishes to access the IPND for the following data items:

- public numbers
- usage codes (business/government/residential/charity/NA)
- type of service (indicates whether service is fixed line or listed mobile)
- all service address data items
- all directory addresses
- service status code (indicates whether service is connected or disconnected)
- pending flag and its removal
- transaction date (date of change in IPND)
- service status date (date of entry on data providers system)
- alternative address flag (indicates service address may not be where service terminates)

The ABS may be collecting unlisted phone numbers from the IPND but would have all fields other than the Public Number and the List Code blanked out in relation to unlisted phone numbers. IIS understands that these records will not have any further use by the ABS and could be removed immediately upon receipt. It is important to note that under Australian Privacy Principle 3 ABS must only collect information that is reasonably necessary for, or directly related to, one or more its

⁸ See Carrier Licence Conditions (Telstra Corporation Limited) Declaration 1997

functions or activities. In future, the ABS is considering requesting authorisation to use the IPND to make contact with survey respondents.

The ABS does currently have Contact Data from the Sensis White Pages, which contains a Residential Telephone Directory and a Business & Government Telephone Directory. The Residential directory contains information including the surname, initial(s), telephone no. and full residential address of individuals; the Business & Government directory contains information of the organisations, including their address, telephone no., website, logo and branches information.

3. Methodology

3.1 Context for the PIA

This PIA is defined by the following parameters agreed between IIS and the ABS:

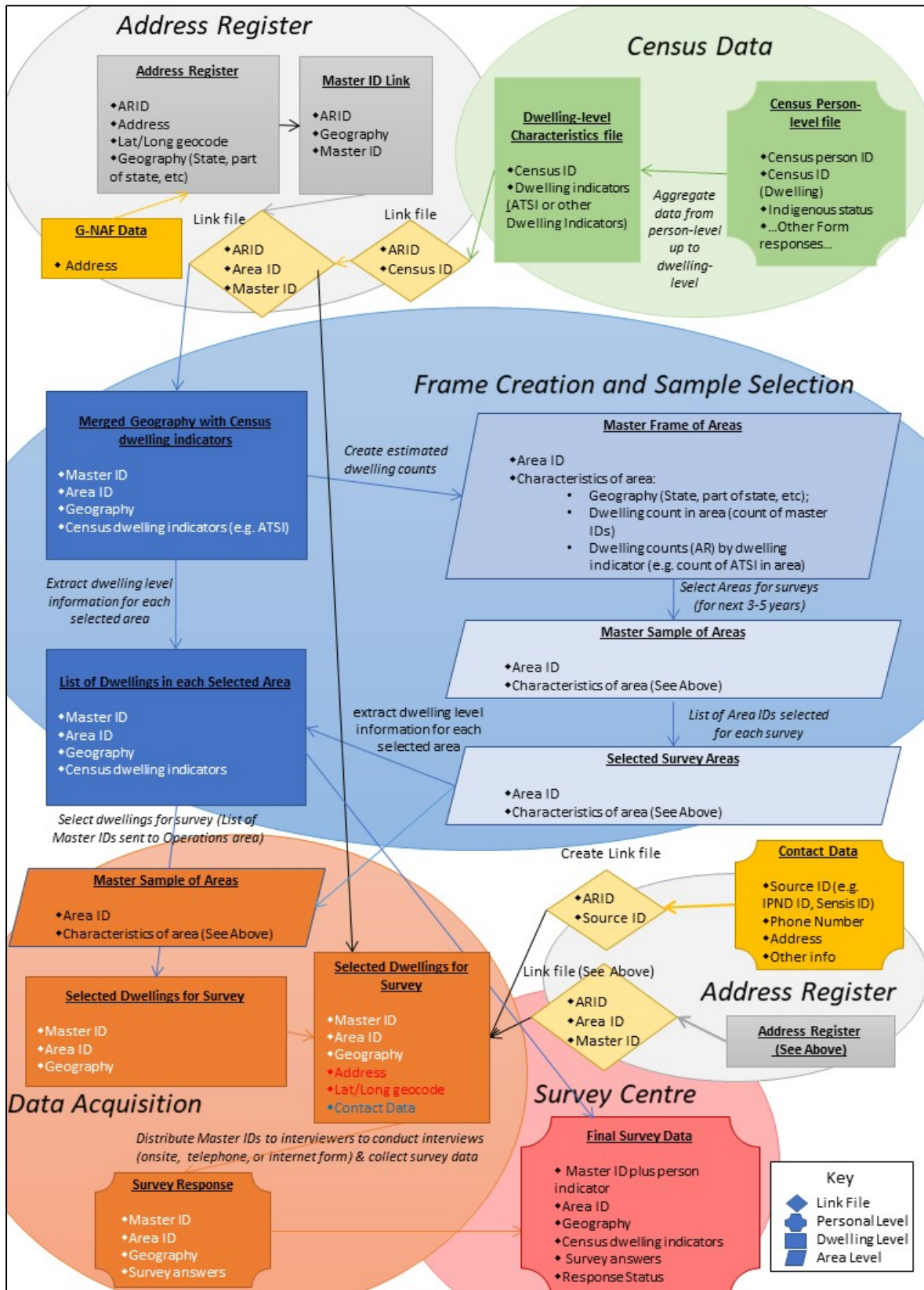
- IIS provides independent and objective privacy advice but not legal advice
- Due to time constraints IIS was not expected to undertake any consultations with community stakeholders
- While IIS carried out sufficient interviews to allow for an assessment of likely risks and gaps, the PIA is not an audit or assurance process

3.2 Process

In completing the PIA, IIS requested and reviewed essential ABS documents to obtain an overall understanding of its two proposals. The documents reviewed are listed in [Appendix 1](#).

On 9 August 2017, IIS conducted a series of on-site, in-depth interviews at the ABS with staff as outlined in [Appendix 2](#), with follow-up discussions on 22 August 2017.

4. Proposed Information Flows



5. Positive Privacy Impacts

In considering the privacy impacts, IIS undertook a high-level analysis of the information flows associated with the proposed linkages. IIS considers the following points have positive privacy impacts or are privacy-friendly practices as they relate to the proposed linkages. IIS notes that other security and access controls are also in place generally within the ABS (though not outlined here):

- Less Contact from the ABS – the first proposal has some positive privacy impact to most people, as they are less likely to be approached (being randomly selected) to conduct a screening/survey if they do not have characteristics of interest to the survey.
- Separation Principle⁹ – IIS understands that no single area in the ABS will be able to see the address and the Census Dwelling Indicator at the same time (see information flow diagram). The Address Register area can only see the address and ARID; the frame creation and sample selection area can only see the master ID, area level geography and the Census Dwelling Indicator; the Data Acquisition area can only see the address and master ID; the survey centre can only see the master ID and survey data (without the address). Access to data is limited on a need to know basis according to role-based access controls. Contact Data Stored Separately – ABS has proposed that phone numbers (such as IPND/Sensis) will not be stored on the Address Register but in a separate folder with access managed by the Director Address Register, which will be released to internal ABS staff on a need to know basis for approved purposes only.
- Concordance files stored securely – ABS has proposed that access to concordance files is tightly managed.
- Retention Period - NATSIHS is not a longitudinal survey and the ABS proposes that linked datasets (which contain addresses and dwelling characteristics) will be deleted three years after the survey in line with section 24 of the *Archives Act 1983* (Cth) and its specific Records Disposal Authority Job No 2001/540 issued in March 2001 by the Australian Archives. However, that same Disposal Authority requires filled in collection forms from respondents to be destroyed when no longer required.
- Structure of Address Register Identifiers – Address Register IDs are sequentially generated rather than generated according to any personal or dwelling attributes, as such do not contain identifiable information.

⁹ See Principle 6, Data Integration Involving Commonwealth Data for Statistical and Research Purposes: Governance and Institutional Arrangements, available at:

[http://www.nss.gov.au/nss/home.nsf/0/00fb7e20e1d56b96ca2577f20016c3db/\\$FILE/data_integration_booklet_5.pdf](http://www.nss.gov.au/nss/home.nsf/0/00fb7e20e1d56b96ca2577f20016c3db/$FILE/data_integration_booklet_5.pdf)

6. Privacy Risks

This Part aims to identify and critically assess key privacy risks created by the proposed linkages. Focus will be on perceived high risk areas. A table which outlines risk of non-compliance with the Privacy Act can be found in [Appendix 3](#), including a risk rating (high, medium, low). However, even if the proposed linkages comply with privacy law, negative privacy impacts may still arise from community perceptions and expectations. IIS considers the following points most relevant to its risk analysis:

6.1 Social Licence and Public Perception

When people trust that their personal information will be used as they have agreed or as set out in statute or a clear privacy notice, and when they accept that enough value will be created, they are likely to be more comfortable with its use.¹⁰ This acceptance is referred to as social licence.

IIS recognises that the potential linkage of Census Data with addresses in the Address Register (obtained from publicly and non-publicly available sources) may be perceived by the public as 'function creep'. Function creep describes a situation where information collected for one purpose is then used for other unintended and unexpected purposes, often beyond the knowledge of the individual concerned. Function creep may erode any social licence the ABS enjoys in its activities.

Function creep relevant in the proposed address-Census Data linkage could appear to include a perceived derogation of the promise made in the 2016 Census privacy policy - *'after data collection and processing, the ABS will remove names and addresses from other personal and household information. Names and addresses will be stored securely and separate from one another. No one working with Census Data will be able to view your personal information (name or address) at the same time as your other Census responses (such as age, sex, occupation, level of education or income).'*¹¹ This is reinforced by the PIA that was conducted in December 2015 on the 'Proposal to Retain Name and Address Information from Responses to the 2016 Census of Population and Housing' which states that "once separated from the Census dataset, addresses would not be brought back together with other personal and household information in the Census dataset."¹²

The ABS has pointed out that under the proposed address-Census Data linkage, addresses are originated from the Address Register, not from the Census Data/Census form. The linkage key is

¹⁰ Defined by the Data Futures Partnership, New Zealand, available at <http://datafutures.co.nz/our-work-2/talking-to-new-zealanders/social-licence/>

¹¹ Census Privacy policy at time of Census (under section Security and retention of personal information, available via waybackmachine at

<https://web.archive.org/web/20160804045458/http://www.abs.gov.au/websitedbs/censushome.nsf/home/privacypolicy>

¹² Australian Bureau of Statistics, 'Privacy Impact Assessment - Proposal to Retain Name and Address Information from Responses to the 2016 Census of Population and Housing', December 2015 p.14, available at [http://www.abs.gov.au/websitedbs/D3310114.nsf/4a256353001af3ed4b2562bb00121564/170fd5a4b684aa3eca257f1e0021a392/\\$FILE/ABS%20Privacy%20Impact%20Assessment%202016%20Census.pdf](http://www.abs.gov.au/websitedbs/D3310114.nsf/4a256353001af3ed4b2562bb00121564/170fd5a4b684aa3eca257f1e0021a392/$FILE/ABS%20Privacy%20Impact%20Assessment%202016%20Census.pdf)

between the Census processing ID and the Address Register ID. As such, the ABS does not plan to derogate from its privacy promise technically as no addresses from the Census forms are linked.

IIS is not in the best position to determine whether Census addresses are technically different from addresses in the Address Register. However, it will point out that the ABS may not be able to convince the public easily that the linkage with addresses in the Address Register (collected from public sources or through ABS canvassing officer observation) is *substantially* different from the linkage with Census addresses (collected from the Census forms). This is especially the case when addresses from the two sources (Census and the Address Register) are largely identical in content and representation, not to mention that Census addresses also play an important role in the linking process.

6.2 Transparency Issues

According to the Australian Community Attitudes to Privacy Survey 2017 conducted by OAIC, contact details and addresses are data points that Australians are most reluctant to provide to business and government.¹³

The ABS issued a statement of intent for exploring the retention of names and addresses from the 2016 Census for amongst other objectives, 'supporting a range of organisational efficiencies, such as the development of an address register, improving sampling, imputation and provider management' in the future.¹⁴ However, to those without an understanding of the meaning of these statistical terms it may not be clear that this would extend to using Census Data to increase the chance of selection of populations of interest in future Household surveys. There is currently very limited publicly available information regarding how ABS could use Census Data for more focused sample design.

Under the Privacy Act, Australian Privacy Principle (APP) 5 requires the ABS to take reasonable steps either to notify the individual of certain matters, or to take reasonable steps to ensure the individual is aware of the fact, circumstances and the purposes of collection.

There is also limited publicly available information regarding the way that ABS collects and uses information about addresses within the Address Register. At this stage (since this is at proposal stage only), there is also no information about how the ABS may collect and use telephone numbers from Sensis White Pages or the IPND. It is worth noting that under APP 5.2(b)(ii), even if the information is not collected directly from individuals, the ABS is obliged to notify of the collection of indirectly solicited personal information into its information collection processes. Therefore APP 5 will apply to addresses transferred from the G-NAF database (if considered to be personal information) and contact numbers and other information from IPND or Sensis White Pages.

¹³ OAIC, Australian Community Attitudes to Privacy Survey, available at <https://www.oaic.gov.au/engage-with-us/community-attitudes/australian-community-attitudes-to-privacy-survey-2017>, p.i, p.4-5

¹⁴ See ABS Statement of Intent – ABS to conduct a Privacy Impact Assessment on retention of names and addresses from responses to the 2016 Census, available at <http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statement%20of%20Intent%20%E2%80%93%20Privacy%20Impact%20Assessment%202016%20Census>

The lack of early notification or ensuring awareness of data collection means individuals are left in the dark and may be caught by surprise anytime with significant potential impact on the ABS social licence. IIS notes that unnoticed or covert collection (e.g. information that is observed, derived or inferred by the ABS without confirming with the individual) is less controlled by individuals and therefore carries higher privacy risks. For example, not only may individuals perceive discrimination (e.g. concerns such as ‘why is the ABS targeting on me?’) or perceive other adverse impacts (e.g. concerns such as ‘will I be made accountable for all activities associated with the number?’), unnoticed or covert collection may also lead to data quality issues as individuals are less likely to be in a position to correct information they are unaware has been collected.

6.3 Opportunity to Consent

From a privacy best practice perspective, it is preferable for an individual to make an informed choice about whether to provide personal information to the information collector. In the 2017 qualitative formative research to inform the development of a marketing and communication strategy conducted by the ABS (the ‘2017 Qualitative Research’), participants indicated that they expected the ability to opt out of the process (both initially at Census data collection time and also at the actual time of the future data collection) if they were uncomfortable.¹⁵

However, for proposal 1, informed choices about the initial data collection in the Census context are not possible because of the mandatory nature of Census – it is against the law for individuals to refuse to complete the Census form. IIS was also advised that it is impractical for ABS to ask Census respondents to consent to their Census Data being used for more focused sampling in the future. This is because consent would lead to the occurrence of sampling bias¹⁶ and greatly affect the quality of the statistics.

For proposal 2, informed choices are not possible because both addresses in the Address Register and contact details from IPND or Sensis are not directly collected from individuals. Rather, they are collected indirectly via third parties (e.g. Australian Communications and Media Authority (ACMA) or PSMA Australia). ABS has sought permission from the ACMA to collect and use the IPND data under the permitted research purposes under the Telecommunications Integrated Public Number Database Scheme 2017.

As mentioned, generally indirect collection of personal information increases privacy risk, as individuals have less control over the accuracy or otherwise of their personal information.

¹⁵ ABS 2017 market research conducted by ORC International (the 2017 Qualitative Research)

¹⁶ Sampling bias refers to a bias in which a sample is collected in such a way that some members of the intended population are less likely to be included than others.

6.4 Consultations on the Proposed Linkages

As discussed above, there is potential for the proposed linkages to be criticised by the community, the media or privacy advocates. Criticisms may also arise because the process of the proposed linkages is misunderstood or misrepresented.

Consultation is a key factor in building social licence to support a project. To date no consultation specific to the two proposals has taken place. While the ABS engaged a Market Research Company to conduct a focus group covering the topic of data linking in 2014,¹⁷ the emphasis of that focus group was on the use of addresses to *improve data integration accuracy in linking, analysing and reporting data*.

Of relevance however, in February 2017, the ABS contracted a market research agency to undertake the 2017 Qualitative Research to inform the development of a marketing and communication strategy (or strategies) for the ABS in 2017. In that research, when participants were asked about the ABS using Census data integrated with other data to better select households for surveys, views on the acceptability of this were mixed. Some participants were very comfortable and felt that such use “made sense” as it would “save resources such as money and time”. However for a few others such activity was “too targeted”, it was “not random” and therefore could be “used to manipulate the outcome.”

Further, a few participants said that they would want to be told upfront that this could potentially happen and why, and that they expected the ability to opt out of the process (both initially at Census data collection time they wanted to nominate if they did not want to be included in future data collections, and also at the actual time of the future data collection) if they were uncomfortable. “You need an option to opt out. It shouldn’t be compulsory. You should be given the choice.”¹⁸

Given that the list containing precise addresses and dwelling information is also of great commercial value (e.g. for target advertisement, for surveillance, etc), and that entities from both the public and private sector are interested in obtaining the list, consultations are essential to address community concerns. Otherwise, the ABS may not be able to provide confidence to stakeholders that security of the information has been maximised and their privacy has been respected.

6.5 Engagement and Consultation with Population Groups of Increased Interest

The communities of increased interest may also voice concerns about the additional risk associated with the proposed linkages. For instance, using NATSIHS as an example, Aboriginal or Torres Strait Islander persons whose dwellings are more likely to be chosen might be worried about the exposure of their identity, or their personal safety if addresses are leaked (which is unlikely given the ABS’

¹⁷ Colmar Brunton focus group, available at

<http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Data+Integration+-+Community+Attitudes>

¹⁸ ABS 2017 market research conducted by ORC International

stringent controls), or having the possibility of being asked to complete more mandatory Household Surveys in the future, or a perception of being monitored or adversely discriminated or studied. They may also find it difficult to dissociate such approaches from other outreach by government agencies at any level of government that they fear will lead to racial discrimination (even though being approached in this way provides an opportunity to be heard and counted). Again, these concerns can be real even if there are stringent controls and safeguards already in place.

This might also lead to a data quality problem, for instance, when some people become less willing to provide answers in Census that will potentially lead to an increased chance of being selected in subsequent surveys.

IIS is aware that round tables have been conducted with respect to NATSIHS in previous years. However, to date they have not specifically addressed the proposed linkages. Rather, as ABS advises the focus has been on tactical and operational issues relating to improving data and increasing participation of Aboriginal or Torres Strait Islander persons in statistical processes.

IIS is also aware that ABS has a NATSIHS Advisory Group that is responsible for providing external input to ABS. However, the purpose of the NATSIHS Advisory Group is to support the ABS in *maintaining relevant survey content and output that meets the needs of NATSIHS users*, rather than to advise on the procedures involved in locating respondents. In practice it is focused on engaging with the community to ensure strong participation in the survey. As such, and as an “Advisory Group, it does not provide any external governance on the proposed approach to sampling of the Aboriginal or Torres Strait Islander Community.

Without adequate and well timed dialogues with the communities of interest, ABS cannot assure itself that the linkage of addresses with census data/contact data align with such community’s expectations.

6.6 Data Security, Retention and Access

Information is both an asset and a liability. The longer the personal information is held by ABS, the more protection, storage and security controls that need to be in place.

The ABS Privacy Policy sets out its personal information handling practices. As an Australian Government agency, the ABS has obligations under the Privacy Act to handle personal information in accordance with the Australian Privacy Principles (APPs). Information collected is protected in accordance with the Australian Government Protective Security Policy Framework and personal information is managed in accordance with the Australian Government records management regime. When no longer required, personal information is destroyed or deleted according to the Administrative Functions Disposal Authority and ABS records authorities (2001/00000540 and 2007/00105946). The Privacy Policy also details how respondents can request access to personal information held about them.

Access to all ABS premises is strictly controlled in accordance with the Protective Security Policy Framework (PSPF). All ABS staff are subject to security checks and are required to sign an Undertaking of Fidelity and Secrecy under section 7 of the Census and Statistics Act – it is an offence under section 19 of that Act for an ABS officer to disclose information obtained under the Act.

As detailed in section 5 separation principles will be followed for these projects. No single area in the ABS will be able to see the address and the Census Dwelling Indicator at the same time. Access to concordance files is limited to a need to know basis according to role-based access controls. IIS is aware that all ABS staff are able to access the Address Register, which increases the probability that in some circumstances it will constitute personal information, for example, if it is linked with survey data.

IIS understands that, tentatively, the list containing all selected dwellings to be surveyed (see Information Flow Diagram) will be compiled nine months before the NATSIHS at the ABS Operation Area. According to ABS staff, this is for the survey design and selection work to project and facilitate the allocation of resources and training of interview staff. Once the survey is completed, the list of addresses will be retained for around three years. This is for analysing and improving future sample selection process, as well as to assist in overlap control – that is to ensure that where possible addresses selected in one survey are not reselected into another survey over a 3 – 5 year period. While these are very legitimate reasons for keeping the information, ABS should not overlook the importance of minimising the duration of the information that it keeps.

IIS also observed that the ABS does not have a specific access policy to the created datasets. APP 12 provides individuals with the right to access their personal information. As such, the ABS is required to provide access to all of an individual's personal information it holds on request, unless an exemption applies. Yet, developing an access policy may be difficult for the created dataset because on the one hand, the information within can be regarded as Statistical Data¹⁹ as part of it (e.g. indicator) is derived from Census Data in which the Return to Source Policy applies; on the other hand other information such as Contact Data or Census dwelling indicators can be regarded as Non-Statistical Data because it is for facilitating the contact of survey respondents. The ABS does not, as such, have a massive database of the personal information of Australians.

Under APP 12, the ABS must be satisfied that a request for personal information is made by the individual concerned, or by another person who is authorised to make a request on their behalf, for example, as a legal guardian or authorised agent. If the ABS gives access to the personal information of another person, this could constitute a disclosure, which may not comply with APP 6. The fact that an individual may not be able to confirm their connection with the dwelling easily makes the policy on access more complicated.

¹⁹ Statistical information is the information provided on the Census form covering the topics prescribed by the Census and Statistics (Census) Regulation 2015, see ABS Census Privacy Policy

7. Recommendations

IIS has made recommendations to address the key risks it has identified in reviewing the proposals for the integration of addresses with the Census Data (Proposal 1) and Contact Data (Proposal 2).

Most recommendations identified are applicable for the integration of addresses with the Census Data (Proposal 1) and Contact Data (Proposal 2). For a recommendation that is only applicable to Proposal 1, a red 'Proposal 1' has been assigned as an indicator on the left of the relevant paragraph; similarly, if any part is applicable only to Proposal 2, a blue 'Proposal 2' will be visible as an indicator on the left of the relevant paragraph.

7.1 Recommendation 1 – Social Licence and Transparency

To retain social licence, it is important that the ABS should not stop at the question 'can the ABS use the data this way', but should go further and ask **should** the ABS use the data in this way? For example, by asking 'what is the public value and benefit in the ABS using the data this way and what is the least possible intrusive means of attaining such public value and benefit?'

IIS considers that should the two proposals be implemented, that ABS should provide maximum transparency to the community on its approach in line with APPs 1 and 5. This includes outlining the consultations that have taken place should the ABS adopt the consultation recommendations in this report. Further, IIS recommends providing information on the ABS website in appropriate locations (such as the Survey Charter and Privacy Policy) including:

- The fact that an Address Register has been created and from what sources of data and what will be stored in the Address Register (ABS can make clear the Address Register alone does not contain identifiable information)
- The policy with respect to the use of the information in the Address Register – including whether other entities will be allowed to use or access the Address Register (if that eventuates)
- The relevant privacy safeguards that surround a more focused sample design and selection approach – a documented protocol for ABS staff to follow for when Unequal Probability sampling techniques are considered appropriate (see Recommendation 4)
- That contact numbers may be obtained from third parties (such as IPND (if this eventuates) and Sensis White Pages), the conditions when they will be used, for example, to contact survey respondents – with an option to opt-out of contact in this manner after first use
- This external PIA or other risk assessment reports that have been done to identify and mitigate privacy risks

Proposal 1

Proposal 2

Proposal 1

In addition, ABS may consider reiterating how the decision for increasing the efficiency of survey operations for certain populations (Aboriginal or Torres Strait Islander persons in this case) was made, such as what were the governance and consultations procedures it has gone through, what

were the major concerns and the mitigating strategies, etc. There should also be adequate opportunities for concerned people to inquire to that effect.

7.2 Recommendation 2 – Consultations with Communities of Increased Interest

Proposal 1 IIS recommends that the ABS should initiate consultations and engagement activities for the communities that will be subject to Unequal Probability sampling (e.g. Aboriginal or Torres Strait Islander persons living in Aboriginal or Torres Strait Islander-sparse areas, persons with a Disability, persons from a Culturally and Linguistically Diverse Background, etc.). The ABS should ensure that its engagement and communications staff have the capabilities necessary to tap into the affected community's sentiment.

Proposal 1 The starting point might be to hold consultations regarding who are the suitable representatives to represent the interests of the community being considered for unequal probability sampling within areas. IIS recommends that the ABS should ensure these representatives have the competence to understand the privacy risks, the objectives of the ABS, the impact on the community of interest and have the credibility to represent that community.

Proposal 1 The form in which these consultations take is dependent on the community of interest. In the case with the Aboriginal or Torres Strait Islander Community, the ABS (e.g. NCATSI) already has strong relationships with key members and representative groups of the Aboriginal or Torres Strait Islander persons, which can be leveraged to determine the appropriate consultation methodology for the specific community.

Proposal 1 IIS also considers it useful for the ABS to obtain feedback on whether the community of interest would be likely to alter their future Census responses if they know they may be used to increase the efficiency by which ABS can select them for other Household Surveys.

7.3 Recommendation 3 – Pilot Testing

Proposal 1 It is already apparent to ABS through its qualitative surveys conducted by ORC International in 2017 that a few people are not comfortable with ABS using Census data integrated with other data to more efficiently select households for surveys. This PIA also identifies the lack of publicly available information which would inform people that use of their Census data is contemplated for more focused sampling. As such, ABS needs to consider whether it should proceed with the first proposal, and if so what steps it will adopt to address concerns in the overall community and communities of interest.

As such, after consultations with the community of interest, and where the ABS is satisfied that respective community sentiment and privacy concerns are addressed and proceeds with the proposals, IIS recommends that the ABS only implement the more focused sampling design methods gradually, starting from a smaller part of the population. A gradual implementation can ensure the ABS will be well aware of the level of comfort people have with using their Census Data for more focused sampling design, and have the time to fine tune the policy, privacy risks or public perceptions that are unforeseen in earlier stages.

IIS also recommends that the ABS run pilot testing to determine the level of effectiveness (including whether people are comfortable with being contacted by phone) as part of the survey. For instance, the ABS can call a small number of respondents using the obtained Contact Data, explain to them how their number is obtained from IPND (if that eventuates) and/or Sensis, and seek their feedback. This would allow the ABS to obtain a view as to the level of sensitivities that may or may not exist around the use of IPND or Sensis data to contact survey respondents. Once testing has been done to determine how well contact in this way is received, adjustments should be made accordingly.

Additionally, in any communications with the community of interest during pilot testing and roll out if proposal 2 proceeds, communications should reflect the level of skill, education and language of the affected population.

7.4 Recommendation 4 – Unequal Sampling Policy

IIS recommends that the ABS develops a clearly articulated policy as to when Dwelling level Census Data will be used for more focused selection of dwellings within areas if it decides to proceed with the Proposal. The policy should list out the prerequisite conditions that would have to be fulfilled before more focused sampling using Census Data at the dwelling level within areas takes place. For example:

- when the respondents are hard to locate (their prevalence is lower than a certain percentage in a particular geographical area)
- when there are no alternatives to locate the respondents (other means have been exhausted but are ineffective to locate respondents with the characteristics of interest)
- when there are legitimate reasons for identifying the respondents (whether reasons are legitimate should be decided by an external oversight body, e.g. the NATSIHS Advisory Group)
- when there are adequate announcements, consultations and engagement prior to the commencement of the process

7.5 Recommendation 5 – Address Register and Address Register IDs

IIS recommends that prior to commencing linkages, the ABS removes Address Register access from staff whose daily operations would not require the use of the Address Register and that Address Register IDs are not contained on survey results. Removal of staff-wide access would reduce the likelihood of the Address Register content being personal information in combination with other data sets.

Minimising access to the Address Register and Address Register IDs is likely to reduce re-identification, through adding data sets together, as Address Register IDs would not be easily linked to addresses if the Address Register is not universally available to all staff.

A unique ID (Master ID) can be used in lieu of the Address Register ID throughout the process, including on the survey frames and final survey datasets. Separate concordance files can be created

to enable address and contact information to be made available for data acquisition, and to enable analysis of Census Dwelling indicators against reported survey data. This would reduce the risk of personal information being revealed to ABS staff involved in the linkages and supports the ABS' strong separation principles.

7.6 Recommendation 6 – Opportunity to Opt-Out of Future Phone Contact

IIS recommends that the individuals contacted by phone should be asked their preferred contact method after ABS staff ring, and if they choose not to be contacted by their IPND-listed public number again, the ABS should ensure that the dwelling should not be contacted with the same IPND-listed public number for that dwelling again in the future.

7.7 Recommendation 7 – Compliance with APPs

IIS recommends that ABS review its compliance with the APPs in light of the risks IIS has identified in the Compliance Table, in particular focusing on the high risk areas relating to open and transparent management of personal information, notification of the collection of personal information, use of personal information and quality of personal information. IIS has provided specific recommendations affecting open and transparent management of personal information as this is likely to have the most impact in mitigating the other risks identified.

1. Appendix 1 - Documents Reviewed

Documents Reviewed

2016 Census inquiry documentation

ABS Announcement and PIA regarding retention of Names and Addresses in 2016 Census

ABS Corporate plan

ABS 2017 market research conducted by ORC International

ABS Organisational structure

ABS Privacy statement

ABS Statement of intent & expectations

ABS Survey Charter

ABS Trust survey, 2010

ABS Trust survey, 2015

Census Form 2016

Census Instruction Letter (Approach)

Census Instruction Letter (Reminder)

Census Privacy Policy, 2016 (at time of Census)

Colmar Brunton focus groups

Data integration focus groups, 2011

Data integration focus groups, 2014

Data Integration principles and protocol

ELG APS Implementation Paper - Address Register (18 April 2016)

Information on ABS engagement with Aboriginal and Torres Strait Islander persons

IPND application form (by the ABS)

MacGibbon review of 2016 eCensus

NATSIHS 2012-13 User Guide

NATSIHS Sampling Methodology presentation

Overview of ABS data integration

Documents Reviewed

Survey participant information pages

2. Appendix 2 - Interviews Held (9 Aug 2017)

Name and Position

Amanda Malekin - Address Register User Interface

Duncan Young, Program Manager - Census 2016 Branch

Edward Szoldra, Assistant Director - Household Survey Methodology

Fiona Blackshaw, Assistant Director - Household and Business Surveys Implementation Team

Fiona Haddon, Assistant Director, National Centre for Aboriginal and Torres Strait Islander Statistics (, NCATSI),

Melissa Gare, Director - Household and Business Surveys Implementation Team

Nick Stathis, A/g Program Manager, Communication and Partnerships Branch

Paul Taylor, Director - Address Register

Stephen Cohen, Assistant Director - Household and Business Surveys Implementation Team

Wolfgang Hertel, Assistant Director - Risk, Planning & Policy Branch Team

3. Appendix 3 - Compliance Risk Table

Australian Privacy Principle	Proposal 1	Proposal 2
<p>APP 1 — Open and transparent management of personal information</p> <p>Ensures that APP entities take steps to comply with the APPs, including having a clearly expressed and up to date APP privacy policy.</p> <p>Risk Level - High</p>	<p>The Census privacy policy does not explain that Census personal data (addresses and auxiliary/attribute information) is going to be linked to support more focused Household Survey sample design and selection</p>	<p>The ABS general privacy policy mentions that contact details are collected from external sources – it would be worthwhile indicating that these external sources include G-NAF database for the Address Registry and IPND/Sensis .</p>
<p>APP 2 — Anonymity and pseudonymity</p> <p>Requires APP entities to give individuals the option of not identifying themselves, or of using a pseudonym. Limited exceptions apply.</p> <p>Risk Level - Low</p>	<p>In some contexts address may be personal information in combination with other data sets</p>	<p>In some contexts contact information may be personal information in combination with other data sets</p>
<p>APP 3 — Collection and means of collecting solicited personal information</p> <p>Outlines when an APP entity can collect personal information and by what means.</p> <p>Risk Level - Low</p>	<p>In some contexts address information may be personal information – it is being collected indirectly for retention in the Address Register, therefore should fit an exception for indirect collection</p>	<p>In some contexts contact information may be personal information – it may be collected indirectly from the IPND or Sensis, therefore should fit an exception for indirect collection</p>
<p>APP 4 — Dealing with unsolicited personal information</p> <p>Outlines how APP entities must deal with unsolicited personal information.</p>	<p>N/A</p>	<p>N/A</p>
<p>APP 5 — Notification of the collection of personal information</p> <p>Outlines when and in what circumstances an APP entity that collects personal information must notify an individual of certain matters.</p> <p>Risk Level - High</p>	<p>Risk that Census Respondents are not aware that the information collected (addresses and auxiliary/attribute information) is going to be linked to support more focused Household Surveys Sample Design and selection (purpose of collection)</p>	<p>Risk that individuals are not aware of the fact that the ABS intends to indirectly collect contact information</p>

Appendix 3 - Compliance Risk Table

<p>APP 6 — Use or disclosure of personal information</p> <p>Outlines the circumstances in which an APP entity may use or disclose personal information that it holds</p> <p>Risk Level - High</p>	<p>Risk that personal information for purpose of Census is being used for another purpose, namely selection of particular households for future surveys – may not be an expected use</p>	<p>Risk that contact information could be used for other purposes. People may be concerned that IPND data is used in this way – may not be an expected use</p>
<p>APP 7 — Direct marketing</p> <p>An organisation may only use or disclose personal information for direct marketing purposes if certain conditions are met.</p>	<p>N/A</p>	<p>N/A</p>
<p>APP 8 — Cross-border disclosure of personal information</p> <p>Outlines the steps an APP entity must take to protect personal information before it is disclosed overseas.</p>	<p>N/A</p>	<p>N/A</p>
<p>APP 9 — Adoption, use or disclosure of government related identifiers</p> <p>Outlines the limited circumstances when an organisation may adopt a government related identifier of an individual as its own identifier, or use or disclose a government related identifier of an individual.</p>	<p>N/A</p>	<p>N/A</p>
<p>APP 10 — Quality of personal information</p> <p>An APP entity must take reasonable steps to ensure the personal information it collects, uses or discloses is accurate, up to date and complete.</p> <p>Risk Level - Medium</p>	<p>The quality of future Census surveys may be compromised if people feel that providing certain data will lead to greater chance of inclusion for future surveys</p>	<p>Risk that contact numbers are not accurate and do not relate to the intended address</p>
<p>APP 11 — Security of personal information</p> <p>An APP entity must take reasonable steps to protect personal information it holds from misuse, interference and loss, and from unauthorised access, modification or disclosure. An entity has obligations to destroy or de-identify personal information in certain circumstances.</p> <p>Risk Level - Low</p>	<p>Risk that the retention period of the datasets is longer than necessary – relevant to the extent the data in question can be considered personal information</p>	
<p>APP 12 — Access to personal information</p>	<p>Risk that access is not provided in circumstances where data constitutes</p>	<p>Risk that access is not provided in circumstances where data constitutes</p>

Appendix 3 - Compliance Risk Table

Risk Level - Low	personal information. Consider interaction with Return to Source Policy	personal information. Consider interaction with Return to Source Policy
APP 13 — Correction of personal information Risk Level - Low	As above	As above



**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