



Submission to the Inquiry into the 2016 Census

Senate Standing Committees on Economics
PO Box 6100
Parliament House
Canberra ACT 2600

21 September 2016

Contributors

This submission was prepared by Andrea Leong, Andrea Finno, James Jansson and Tom Gordon on behalf of the Science Party.

Contact details

Email: secretary@scienceparty.org.au

Introduction

The Science Party has a number of concerns with the 2016 Census, some of which we expressed on our blog in the lead-up to and wake of census night.

Our two greatest areas of concern are the increased potential for misuse of census data, and the risk of tainting the dataset due to individuals' concerns about the former.

The Science Party is a strong advocate of digital and offline rights including the right to privacy. The ABS has a strong track record of protecting the public's information, in line with both legislation and community expectations. Its response to privacy concerns surrounding the 2016 Census has been to reassure the public that appropriate checks and balances are in place to minimise the risk of misuse, but actual details have been scant and have put few minds at ease.

The Science Party holds open and efficient government as one of its core principles. The lack of communication from the ABS regarding both privacy concerns and the failure of the census website on the 9th of August undermine public trust in the motives and competence of the organisation. The Science Party urges the ABS to uphold integrity above all, as their work allows the best possible planning for our future. The public's confidence in the ABS must not be shaken. It is far better to gather some of the available data and be confident that it is accurate than to use more powerful methods of analysis but be unsure about the integrity of census data.

Given the issues surrounding the 2016 Census, the Science Party urges the ABS to not pursue fines for failure to complete this census.

A. Preparation, administration and management of the Census

The first item of concern regarding the 2016 Census was the lack of transparency surrounding the decision to retain names and addresses for longer than previously, and to match census data to other datasets (privacy concerns are discussed in more detail in section F). The ABS announced on 11 November 2015¹ that it would undertake a Privacy Impact Assessment (PIA), with feedback accepted until 2 December. The Science Party notes that the consultation period of three weeks was short and poorly publicised; and that the PIA was conducted internally, contrary to best practice.

The ABS announced on 18 December 2015² that names and addresses would be retained; no period was specified, implying indefinite retention of these identifiers. This ambiguity drew complaints, after which the ABS specified a retention period of up to four years³. This concern

¹<http://www.abs.gov.au/websitedbs/censushome.nsf/4a256353001af3ed4b2562bb00121564/2986cb17314b5d02ca257efa000b839b>

²<http://www.abs.gov.au/websitedbs/censushome.nsf/4a256353001af3ed4b2562bb00121564/cb602f87b5ba4648ca257f22007a7bea>

³<http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Retention+of+names+and+addresses+collected>

should have been identified by the PIA, and the initial omission of this information appears as either a lack of planning or backtracking in the face of public disapproval.

The eCensus initiative

Online completion assists in the goal of timely and accurate handling of census data. However, if the system in place is not able to handle the process, then the result, as we have seen, is less data than would have otherwise been collected. The Science Party advocates the uptake of technology where it is safe and advantageous to do so, and the 2016 eCensus appears to have fulfilled neither of these criteria.

B. Data handling

The scope of the data collected in the 2016 Census was no different to previous years, and the Science Party understands that collecting names may increase the accuracy of responses (according to the ABS) as well as allows the ABS to determine who has not completed a Census.

While the potential for misuse of census data is just that: potential, future and unproven, it was well-known by census night that some Australians were wary of providing their names and/or addresses in the light of new, longer retention periods and dataset matching. It would have been a simple matter to allow respondents to opt in to data linking as for the Time Capsule question; while this would likely cause certain populations to self-select for data linkage, it would uphold the prime objective of collecting accurate data.

Communication of how data were to be handled

Perhaps those who were concerned about data matching were in the minority. However, those who were fully informed about data matching and its implications were also in the minority. The ABS in its PIA noted that focus group participants “indicated a general level of support for retaining names and addresses⁴.” However, not mentioned in the PIA were a few participants’ significant concerns about identity theft or other forms of misuse; nor the finding that not a single participant was aware of the data-linking activities of the Australian Census Longitudinal Dataset (which has so far linked responses the 2006 and 2011 Censuses); nor that participants generally felt that the data-linking activities of the ABS were not advertised widely enough and should be noted on the front page of the census form⁵.

C. IT arrangements

It seems almost negligent of IBM to have accepted a contract to create a one-off system built to handle no fewer than several million online forms in an evening. Infrastructure to handle this volume of traffic (e.g. large social media sites) is typically built up slowly and proven over time.

⁴[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)

⁵[http://www.abs.gov.au/websitedbs/D3310114.nsf/4a256353001af3ed4b2562bb00121564/82b049b9ee89038eca257bac0017f7f7/\\$FILE/2014%20Report_Colmar%20Brunton.pdf](http://www.abs.gov.au/websitedbs/D3310114.nsf/4a256353001af3ed4b2562bb00121564/82b049b9ee89038eca257bac0017f7f7/$FILE/2014%20Report_Colmar%20Brunton.pdf)

It is also concerning that the ABS was rated as being 'internally secure' but not 'cyber secure' by the Australian National Audit Office at the bureau's most recent audit in 2014⁶. The latter rating, which denotes resilience against external cyber attacks, seems as though it should be a prerequisite for moving from a paper-based to an online census.

D. Website shutdown

While it is disappointing that the eCensus website was unable to continue to operate throughout census night, the shutdown was an appropriate precaution in response to an apparent DDoS attack (although to the best of our knowledge no evidence for such an attack has been made public, and the same outcome on census night would have occurred if the website was simply unable to handle the amount of traffic caused by Australians legitimately attempting to log in to complete their forms as directed).

E. Response rate

Certainly many Australians who attempted to complete the census online, but were unable, simply have not yet gotten around to completing a form. Others who requested a paper form will have delayed completing the form, waiting for communication from the ABS as to whether the data will be useable.

Anecdotes suggest that many Australians discarded the letter from the ABS in the weeks prior to the Census as they thought it was only an unimportant reminder, then were unable to get through on the phone to request a replacement code—if indeed they realised their error. It is very possible that this affected certain demographics more than others.

Other accounts include stories of being harassed and given incorrect information by census collectors, leading to frustration and double submission⁷. Ultimately this undermines the confidence the public once had in the ABS, through both direct negative experiences and the knowledge that the 2016 census data has been compromised.

Mitigating a low response rate

We would expect that the ABS will have rigorous statistical models that predict the accuracy of data for given numbers of respondents, as well as predictions based on sound reasoning about the integrity of the submitted answers.

F. Privacy concerns

It has been policy in previous years to destroy all identifying data from census forms within 18 months, and only used to ensure that a person had completed a census. However, Australian Statistician David Kalisch made the shocking admission that our data has been used in ways that we were assured was not the case: "*They've done it under the guise of: 'this is while we are*

⁶<https://www.anao.gov.au/work/performance-audit/cyber-attacks-securing-agencies-ict-systems>

⁷https://www.reddit.com/r/australia/comments/539pqm/zero_tolerance_abs_threatens_two_million/d7rj7ix

processing the data' [...] They've done linkages, they've done other things. What's happening now is we are being more transparent about it."⁸ Perhaps Kalisch was referring to the longitudinal matching which began in 2006, using 5% of census responses without the respondents' consent or knowledge.

The ABS has generally enjoyed a very high level of trust from the Australian public, but confidence in the ability—or perhaps the will—of the ABS to keep data secure is understandably shaken in the wake of this admission and following the events of the 2016 Census.

Data linking

Data linking—both longitudinally from census to census and between diverse datasets—is a powerful tool. However, this can only be achieved by the use of a repeatable unique identifier, which creates new risks. Previously, a breach of privacy surrounding census data would expose individuals to having only their most recent census responses known. Now, malicious re-identification of anonymised data could see the release of all of an individual's demographic data and health and financial records over time.

The implications of allowing other government departments access to census data must also be properly addressed. Among the things we don't know is which bodies will be granted access to census data. Among the things we do know is that the 'Guidelines on Data Matching in Australian Government Administration' are entirely voluntary.

Statistical Linkage Key generation

Information about how the Statistical Linkage Key would be generated was hard to find, forcing an FOI that resulted in essentially a shoulder shrug⁹. This lack of transparency—or worse, lack of competence—understandably further undermined the public's trust in the ability of the ABS to keep data confidential.

G. Purpose, scope, regularity and cost and benefits

The Science Party loves good data and is a strong supporter of the aims of the Australian census, and supports maintaining a five-yearly census. This is precisely why we were so opposed to the new policy of mandatory data linking, which might lead some respondents to omit at least some answers or give inaccurate ones to protect their privacy regarding questions that they felt were too intrusive.

It's bad science to collect data while knowing that the data will be compromised. In science, if the results are not good, either the hypothesis or experimental method must be changed. In this case, many and varied sources asserted that the method (census data handling) would taint the results, and yet the method was not changed to accommodate these concerns, so we will have bad results. This refusal to adjust to available evidence shows a lack of scientific literacy.

⁸<http://www.smh.com.au/federal-politics/political-news/census-the-abs-has-been-quietly-holding-on-to-our-names-20160721-gqax2e.html>

⁹https://www.righttoknow.org.au/request/method_for_creation_of_linkage_k

H. ABS resourcing

The Science Party is sympathetic to the fact that the ABS is under pressure to produce high-quality findings despite a shrinking budget and staff losses, with the 2016 Census in its entirety on the chopping block at one stage. We feel that this is typical of a government that wants to 'destroy the evidence' (having recently defunded the Bettering the Evaluation and Care of Health program, which studied GP visits, and the Office for Learning and Teaching, which aimed to improve higher education delivery).

Any quick savings made by abolishing these offices or skipping a census are a false economy as we will not only fail to gather metrics, but also lose expertise, which would have allowed the most efficient delivery of the most appropriate infrastructure.

The Science Party urges the Australian government to resource the ABS at a sustainable level that allows the bureau to continue to collect and analyse useful datasets without having to compromise its principles.

I. Ministerial oversight and responsibility

J. Other matters

Reversion to paper forms

Many Australians have received a paper copy of the Census without having asked for it (or received a second, and in some cases a third, copy after requesting a paper form). This seems like an admission that the eCensus has failed. Before attempting to push the majority of census responses online in future years, a full explanation of why a future eCensus would not suffer from the same issues is to be expected.

Fines for failure to complete the 2016 Census

There has been a conspicuous lack of communication regarding whether financial penalties for refusing to complete the 2016 Census will be pursued (although the government and opposition have agreed that fines will not be suspended by an act of parliament). Given the confusion and errors surrounding the 2016 Census process, it is our recommendation that these fines not be pursued, as a show of good faith.