

# **2006 Census of Population and Housing - Australia Innovations to Meet the Challenges**

**Paper Prepared for 22nd Population Census Conference, Seattle, March 7-9, 2005  
Paul Lowe, Director, Population Census Program  
Australian Bureau of Statistics (ABS)**

## **1. Introduction**

The next Census of Population and Housing in Australia will be held in August 2006. Strategic directions for the next Census have established seven key goals that provide the focus for planning and development:

1. Ensure the relevance of the data to users and providers;
2. Maintain provider load at a similar level as the 2001 Census;
3. Revise field procedures to incorporate internet return of census questionnaires and cost effective collection in difficult to enumerate areas;
4. Improve the timeliness of the release 2006 Census data compared with 2001 Census;
5. Maintain at least the same high level of 2001 Census data accuracy;
6. Reduce the real per capita cost of conducting the Census;
7. Maintain Census privacy and the public perception of privacy.

There are many challenges facing the ABS in meeting the seven goals mentioned above. This paper provides a brief description of some of the innovations the ABS is planning to adopt to meet these challenges.

## **2. Topics**

At the present time the ABS is awaiting Government consideration of the topics to be included in the 2006 Census. Attachment A reflects the proposal that has been put forward. The key changes proposed compared to the 2001 Census being:

- the inclusion of 'number of children ever born' - this was previously collected in 1996 and is required for estimating future fertility of sub-groups of the population and the population as a whole for the purpose of constructing population projections;
- the inclusion of a 'need for assistance indicator' i.e. disability - this will focus on the need for assistance with activities such as self care, mobility and communications;
- the inclusion of an 'unpaid care' topic - while there was demand for inclusion of an unpaid work topic it was concluded that unpaid care was the highest priority and the most amenable topic for collection via a census vehicle. The inclusion of this topic will require a page to be added to the census questionnaire for which the ABS will be seeking extra funding from Government; and
- the modification and reduction of the 'information and communication technology' topic - the 2001 Census addressed computer and internet use by individuals. For 2006 it is proposed that these questions be replaced with a household level question on access to the internet.

## **3. Electronic Census Questionnaire (via Internet) and Field Communications**

The use of an electronic Census questionnaire has been under investigation since March 2002. Internally developed facilities have been used in recent tests and an industry partner (IBM) has recently been selected for the development and hosting of an electronic option for the 2006 Census. A facility to be used in the 2005 Dress Rehearsal is currently being developed in conjunction with IBM.

Recent testing has shown that only around 5%-10% of the population may choose to use this option. The focus of the testing has been on both understanding the use/respondent experience of the questionnaire through useability testing and exploring what changes may be needed to existing field processes to support the introduction of this questionnaire.

Evaluation has also started on the statistical impact of multi-modal methods for completing and returning of census questionnaires. As the electronic questionnaire will either replace or be used in conjunction with the paper questionnaire, there will be situations of duplicate response for both dwelling and personal information. With the recent tests showing 0.5% of all dwellings returning both paper and electronic questionnaires, this issue does need to be understood and procedures developed.

Providing an electronic questionnaire also demands that suitable levels of accessibility are included to ensure that population groups such as those with vision impairment are able to use the facility. National Information Library Service, a joint project by the Royal Blind Society and Vision Australia, has been providing assistance and working closely with ABS as this project develops. Blind Citizens Australia is also being consulted, as are the Office of Disability, as plans on the strategy surrounding the electronic questionnaire are developed. Testing will include providing a telephone based fulfillment method for people with vision impairment who do not receive assistance from their Collector.

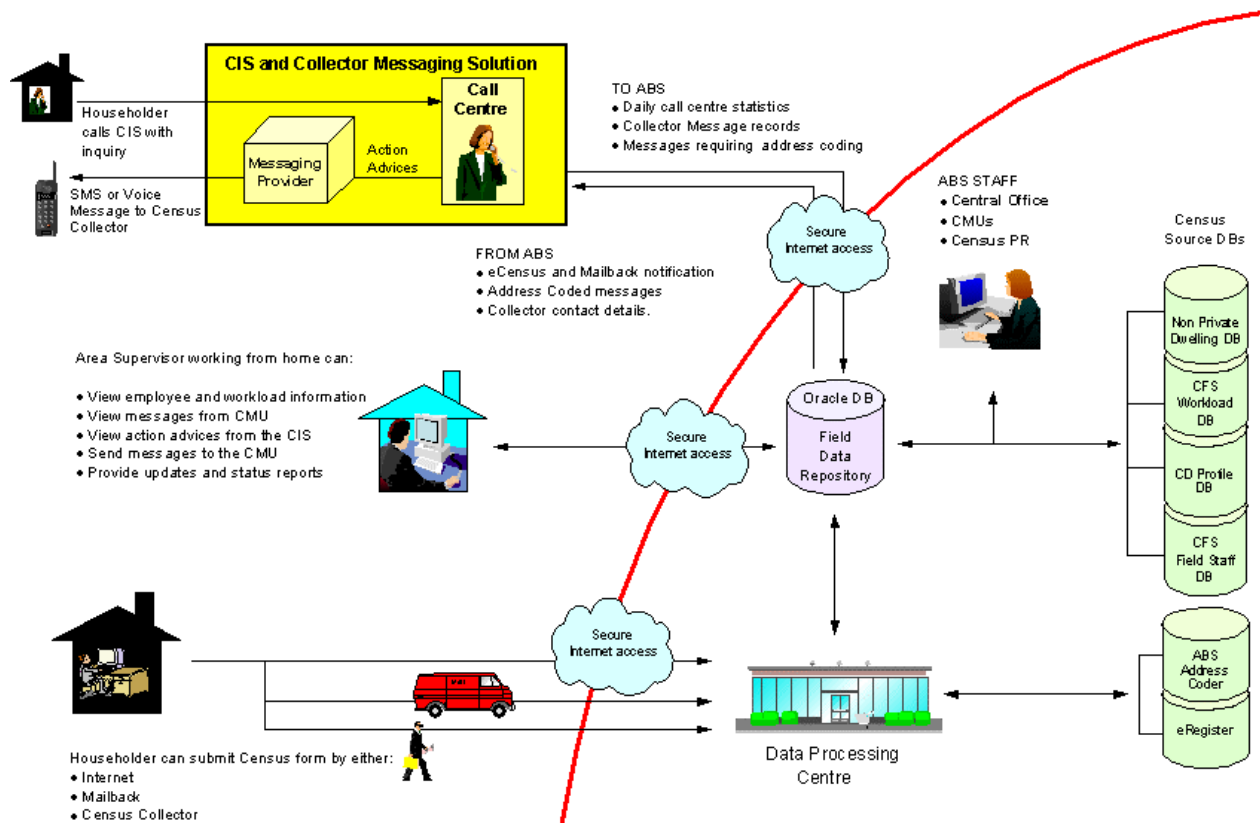
An electronic Census questionnaire option requires field communication processes be modified to inform collection staff of completed electronic census questionnaires being received. The use of Short Message Service (SMS) messages being sent directly to collector's mobile phones has been successful in recent tests. It is anticipated that all relevant actions and notifications related to questionnaire receipt (mail-back or electronic) as well as requests for assistance, and similar, received by the Inquiry Service will be sent directly to collector's mobile phones. From a recent survey, over 70% of collectors own and use a mobile phone, and for those that don't, the messages will be sent to their home phone message bank.

Area Supervisors, managing around 10 collectors each, will have access to an Internet based management information facility which will provide copies of the SMS messages going to collectors and other management information. A recent survey of 2001 Supervisors showed that over 80% have, and use, Internet in their home, and previous techniques around information from District Coordinators or via fax or similar will support the remaining.

For 2001, it was the Area Supervisors role to act as a conduit for messages between the Census Inquiry Service and collectors. As they will no longer be undertaking this activity directly, the expectation being built into procedures is that the Area Supervisor will spend more time in the field.

The following diagram shows the proposed field communication processes.

Diagram 1: Census Field Communications



#### 4. Enumeration

##### Follow up of Refusals and Non-Contacts

Australia has only ever followed up in a consistent fashion outright refusals - i.e. those people who state to the collector that they are not going to complete the census questionnaire. Last census there were only 7,000 households in this category, over half of which completed a questionnaire once the Area Supervisor had visited them. Procedures for this category are being improved with earlier notification and follow up being undertaken. In recent tests, a follow-up letter was sent to every household with another census questionnaire. The householder was requested either to complete and return the census questionnaire or, if they were not at the dwelling on census night, to provide their census night address. A number of the letters were returned indicating an alternative address and that the dwelling was unoccupied on census night.

A much larger category are those households where the collector does not make contact on collection or the householder states that they are going to mail a

questionnaire back but never do so. In previous censuses, the only widely advertised option for returning a census questionnaire was to the collector. There was no effective follow-up of the households that opted to mail a questionnaire back. The Internet option will be widely advertised and provides an easy way-out (i.e. the householder tells the collector that they will complete an Internet questionnaire to get rid of the collector but fails to do so).

Unlike previous censuses, SMS will allow timely feedback to collectors on those households for which a questionnaire has not been received. While planning is in its early stages, it is likely that any households that have not provided a questionnaire by the end of the collection period will be provided with a new blank questionnaire, and will also receive a letter delivered by the collector stating the requirement for a completed questionnaire to be supplied.

### *Harder to enumerate areas*

The 2001 Census saw a rise in non contact dwellings particularly in secured apartment buildings and non private dwellings (e.g. hotels). Extensive testing has been conducted over the past two years targeting these difficult to enumerate areas and a proposed strategy has been developed.

The main focus of the strategy will be establishing a discrete team within each Census Management Unit which are located in each State in Australia. This team will have prime responsibility for implementing the strategy for secure apartment buildings and non-private dwellings within the central business district areas of major metropolitan areas. Key features of this strategy, which is still under development, include:

- early identification of relevant buildings, including access methods into, and contacts for, secure apartment buildings,
- small teams (swat style) of collectors working together to enumerate buildings which do not have an on-site management structure,
- use of on-site management structures,
- focused PR for these buildings, including posters in car parks and lifts, and
- encouraging the use of the internet and mail back as a first option for return of the census questionnaire;

## **5. Processing**

The basis of processing in 2001 was to scan the paper questionnaires and process the resultant images through Intelligent Character Recognition Systems and automatic coding systems. For the 2006 Census the emphasis is to build on these systems and improve efficiencies and quality of data. Processing of both the electronic and mail back questionnaires are also being integrated into main processing.

A prototype of the processing system has been developed and implemented for recent tests. Two major changes have been made to the architecture of the system: implementation of a new data model that utilises Oracle; and, workload management that allows teams of specialised coders to code one or more topics, rather than all topics for a collection district. Early testing of the system has shown that these changes will realise productivity gains whilst also improving data quality.

A new imputation methodology is also being implemented for 2006. This method is a nearest neighbour hot decking approach that uses dwelling structure and credible source information on number of persons that is obtained from collectors in the field. Testing of this algorithm on 2001 data has shown that this new approach reduces the problems experienced in 2001 with the over imputation of persons in certain dwelling structures (e.g. flats and units). Hot decking will also be used for imputing age, sex, usual residence and marital status.

## **6. Dissemination**

The 2006 Census Output Strategy is currently being developed. The following strategies are currently proposed:

- place of usual residence as the basis for dissemination of standard census tables;
- development of an expanded range of products and services based around the ABS Internet website;
- rationalisation of the range and complexity of CD-ROM products; and
- timely release of data through a two-phase release strategy with an increased number of first release data items.

### *Place of usual residence*

Australian censuses have traditionally given priority to place of enumeration data (the census night population) as this could be processed quickly and was the only data available at the collection district (enumeration area) level. For 2006 Census, place of usual residence data will be available at the collection district level.

### *Products and services based around the ABS Internet website*

The focus of standard releases of census data for both 1996 and 2001 Censuses has been community profile data - a number of one and two dimensional tables available for small geographic areas, down to and including collection districts. In 2001 Basic Community profiles down to the statistical local area were made available free of charge on the ABS website. For the 2006 Census, a similar range of small area census data will continue to be available as for the 2001 Census, however it is intended that there will be a wider range of data in the form of data cubes available at the website that can be tailored to meet individual needs.

The ABS is exploring the provision of an on-line tabulation service that will provide appropriate products that meet the needs of different types of users of census data. This service will provide a wider range of census data as standard products than for the previous two censuses and improve the useability of the census data

Electronic dissemination via the internet is seen as the means of accomplishing this. As well as delivering an expanded range of data from the 2006 Census, electronic dissemination allows the data delivery to be tailored to the level of sophistication of the client and increasing the ability of the client help themselves.

In line with this direction the number of printed publications will be reduced and most CD ROM products will become web based.

The aim is to produce functionality to satisfy the different levels of users. For the most part, the aim is for this functionality to be seamless from simple to sophisticated with the clients being lead by the nature of the query or analysis they are wishing to undertake. The aim is also to develop functionality that can be used with a wide range of non-census data - particularly for small area data. As well as community profile type tables with the focus on geographic disaggregation, it is proposed also to produce a range of data cubes with the focus on various population groups or "topics".

#### Users of census data

Users of census data can be grouped in the following manner:

##### *Novice/Infrequent Users*

- Majority of public users
- Limited number of decisions/choices user has to make
- Low level of complexity, i.e. fixed tables
- Ease of use essential - accessibility over functionality

##### *Intermediate/Frequent Users*

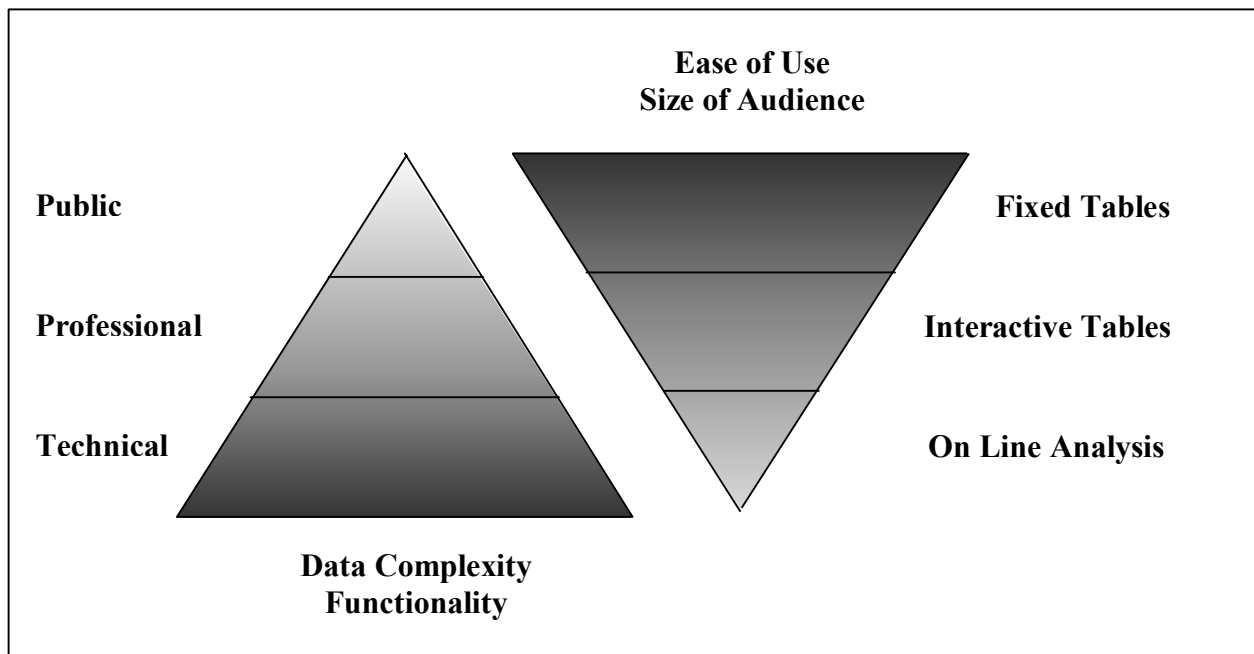
- Subscribers
- Medium level of complexity i.e. interactive tables
- Willing to sacrifice some ease of use for power

- Power needs to be controllable
- Choices/decisions easy to make

*Expert Users*

- Technical
- Limited audience
- On line analysis
- Power is King – want to be able to do everything.

The diagram below maps the client groups against the nature of the products being proposed:



Product differentiation will be used to assist clients in making their choice. The service:

- will provide simple and intuitive web based data and area selection for the three user types and six census products.
- have a main focus on the small area data product.
- provide a flexible tabulation offering table manipulation, charting, mapping, relationship analysis and export capabilities.

The following table shows the range of products that will be available.

Census Product	User level	Concept	Main attributes	Data
Quick stats	Novice	Quick and easy access to basic data for an area of interest	<ul style="list-style-type: none"> <li>- Predefined areas (known to the general public)</li> <li>- Predefined topics</li> <li>- Pick and area</li> <li>- Here's your data</li> <li>- Easy bench marking</li> <li>- Quick in, quick out</li> </ul>	<ul style="list-style-type: none"> <li>- Areas such as State, City, Suburb or Post code</li> <li>- Classification counts</li> </ul>
Community Profiles	Novice	Community Profiles outputted as xls files. Cost \$10 per file for areas below SSD.	<ul style="list-style-type: none"> <li>- User choose an area</li> <li>- User chooses a profile</li> <li>- The may have to pay for the profile</li> <li>- The profile is downloaded</li> </ul>	<ul style="list-style-type: none"> <li>- Rich area set including CCD</li> <li>- Standard profile xls downloads</li> </ul>
Topics	Novice / Intermediate	Primary interest in a specific topic which dictates geographic access level. Only select one area	<ul style="list-style-type: none"> <li>- For users who are interested in a particular subject</li> <li>- Start off via topic selection</li> <li>- Various topic search methods – such as key word (eg Lebanese) or topic based hierarchy</li> <li>- Topic selection will restrict geographic selection</li> </ul>	<ul style="list-style-type: none"> <li>- All areas in the geographic classification</li> <li>- cubes (CP's)</li> </ul>
Community data	Novice / Intermediate	Primary interest in an area which dictates topic availability. Only select one area	<ul style="list-style-type: none"> <li>- For user who are interested in pre defined areas</li> <li>- Area selection will occur via key word searches, geographic hierarchies and maps</li> <li>- The area selected with determine the data available</li> </ul>	<ul style="list-style-type: none"> <li>- All areas in the geographic classification</li> <li>- cubes (CP's)</li> </ul>
Detail area data	Intermediate	Low level geographic data interest requiring easy map based selection	<ul style="list-style-type: none"> <li>- Similar to CDATA – thematic mapping of many areas</li> <li>- Users can define their own areas</li> <li>- Geographic level will dictate data available</li> </ul>	<ul style="list-style-type: none"> <li>- All areas in the geographic classification</li> <li>- cubes (CP's)</li> </ul>
Table builder	Expert	Ability to build your own tables from the full census database	<ul style="list-style-type: none"> <li>- No restrictions on data selection</li> <li>- Enforced random rounding on results</li> <li>- User will be expected to understand basic Census concepts (what to count etc)</li> </ul>	<ul style="list-style-type: none"> <li>- 2006 Census master file</li> <li>- Full geography</li> </ul>

### *Timely release of data*

The ABS will continue the two stage release of data as for the two previous censuses. First release data, which will cover eighty percent of census topics, will be released in June 2007, eleven months after census night. Data for the remaining topics will be released by November 2007.

## **7. Geography**

### *Geocoded National Address File*

In association with key partners, the AEC, CentreLink and Australia Post, the Geography Program has facilitated the development by PSMA Australia of the Geocoded National Address File (G-NAF). This was released on 2 March 2004. G-NAF is the first authoritative database of addresses for all of Australia. It has received support from the State and Commonwealth mapping agencies which jointly own PSMA Australia. It will be updated quarterly and is very quickly establishing itself as the definitive geocoding reference file for Australia.

### *Mesh Blocks*

Across its statistical collections the ABS uses a common standard for geographical areas known as the Australian Standard Geographical Classification (ASGC). Local Government Area is a key spatial unit of the ASGC. Within an LGA, the smallest building block of the ASGC is the Census Collection District (or enumeration area), typically comprising 100 to 220 households. CDs are commonly used for detailed spatial analysis of Census data at the local level. CDs are however subject to certain constraints and consequently they do not always coincide, in shape and size, with areas of interest such as neighbourhoods, electoral boundaries and various catchment area boundaries. This in turn means that Census data can only be approximated for these areas.

The ABS proposes to overcome this shortcoming by developing a new micro-level geographical unit known as a Mesh Block. Mesh Blocks will contain a minimum of 20 to 50 households (about one fifth the size of a CD) and will be created to align with a wide range of administrative and natural boundaries. After the 2006 Census very basic census data will be available at the Mesh Block level, perhaps only number of dwellings and population counts, but a range of census data will be available for combinations of Mesh Blocks to meet clients' individual needs.

Once designed and readily available, Mesh Blocks have the potential to become a new building block of Australian geography so that a government organisation, business or Local Government developing a set of regions for capacity planning, community development, service delivery or indeed any form of regionalisation, would build up its areas simply by adding Mesh Blocks together in whatever combination best suited its purpose. This in turn would ensure that census data could be accurately derived for those regions. For Mesh Blocks to fulfill their potential, however, it is vital that State and Local Governments inform the ABS on how to define Mesh Blocks in a way that would maximise their use not only for obvious statistical purposes but also as the building blocks for the other various zones and

regions which are not defined for statistical purposes but for which statistical data is eventually needed.

## 8. Linking of 2006 Census of Population and Housing data

The 2006 Population Census provides an opportunity to establish the basis of a longitudinal data set from successive censuses, and/or to link census and other ABS data sets. This proposal was canvassed in the recent public consultation on 2006 Census contents and procedures.

Public reaction to the following three options is being tested:

- linking of census data with data collected in successive censuses, other ABS run collections and with Births and Death data,
- as above, plus linking of census data with health data;
- as above, plus linking of census data with other administrative datasets brought into the ABS under the *Census and Statistics Act*.

A two-stage process of focus group testing has been undertaken to assess possible public reaction to data linking. The key messages to date have been:

- a) About one in ten of the focus group participants have concerns about data linking. These concerns were to some extent allayed if the purpose of data linking was clearly explained; data linking was restricted to data sets that the ABS was already perceived to have (e.g. births and deaths registrations and other ABS run collections – first option above); and assurances that identifiable data would always remain within the ABS and could not be accessed by any other organisation.
- b) There was a significant polarisation of opinion once it was suggested that data linking be extended to other "non ABS" data (second and third options above). In particular, for some participants, linking with health data (such as hospital morbidity) was totally unacceptable, to the extent that they would either not complete the census questionnaire or not answer the census questions correctly. In some of the focus groups, those opposed to data linking on this point were able to persuade fence sitters or supporters of data linking to change their opinions.
- c) ABS has a high reputation in the community. There was a great deal of trust that ABS would protect the confidentiality of the information provided to it. Most focus group participants had reservations about, or did not understand, the role of the Privacy Commission or ethics committees.

ABS is currently compiling information on possible uses of linked data sets that will assist in further public consultation and as the basis of a possible new policy proposal seeking additional funding from Government.

ABS is currently developing a communication plan for wider public consultation on data linking to be undertaken early next year. No decision has been made on whether to go ahead with data linking. The final decision will depend on an assessment that community reaction to data linking will not impact adversely on the quality of the census and funding is forthcoming from Government. A final decision will be made in May 2005.

**PROPOSED TOPICS FOR THE 2006 CENSUS OF POPULATION AND HOUSING**

Name	Attendance at an educational institution
Age	Type of institution being attended
Sex	
Marital status	Highest level of schooling
	Education qualifications
Relationship (family structure, including persons temporarily absent)	Level of qualifications
	Field of qualifications
	Income
Address at Census time	
Usual residence at Census time	Labour force status
- One year ago	Occupation
- Five years ago	Industry
	Hours worked
Country of Birth	Unpaid Care
Whether parents born in Australia	Journey to work (destination zone coding)
Ancestry	Mode of travel to work
Year of arrival	Number of motor vehicles garaged
Citizenship	
Aboriginal Torres Strait Islander origin	Number of bedrooms
Language spoken at home	Tenure type
Proficiency in English	Rent
Religion	Housing loan repayment
	Household Access to the Internet
Need for Assistance Indicator	
	Private dwelling structure (classification)
Number of Children Ever Born	Non-private dwelling (classification)