UNITED NATIONS
STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEN COMMUNITIES
(EUROSTAT)

Work Session on Electronic Raw Data Reporting
(Geneva, 6-8 November 2006)

The Australian 2006 Census and the Internet

Paul Williams, Australian Bureau of Statistics

 Summary

1. The role of the Australian Census of Population and Housing is to accurately measure the characteristics of all people in Australian territories on census night. The census aims to provide high quality data for small geographic areas and for small population groups.

2. This paper briefly outlines the processes that were put in place to provide an Internet option to the Australian public at the 2006 Census and the experience with the solution.

3. The use of the Internet provided new opportunities to reach out to geographic areas and population groups that were difficult to enumerate using traditional census methodology. As well there are a range of risks that needed to be managed to ensure that the implementation of the Internet solution was both operationally sound and secure and the quality of the census was maintained.
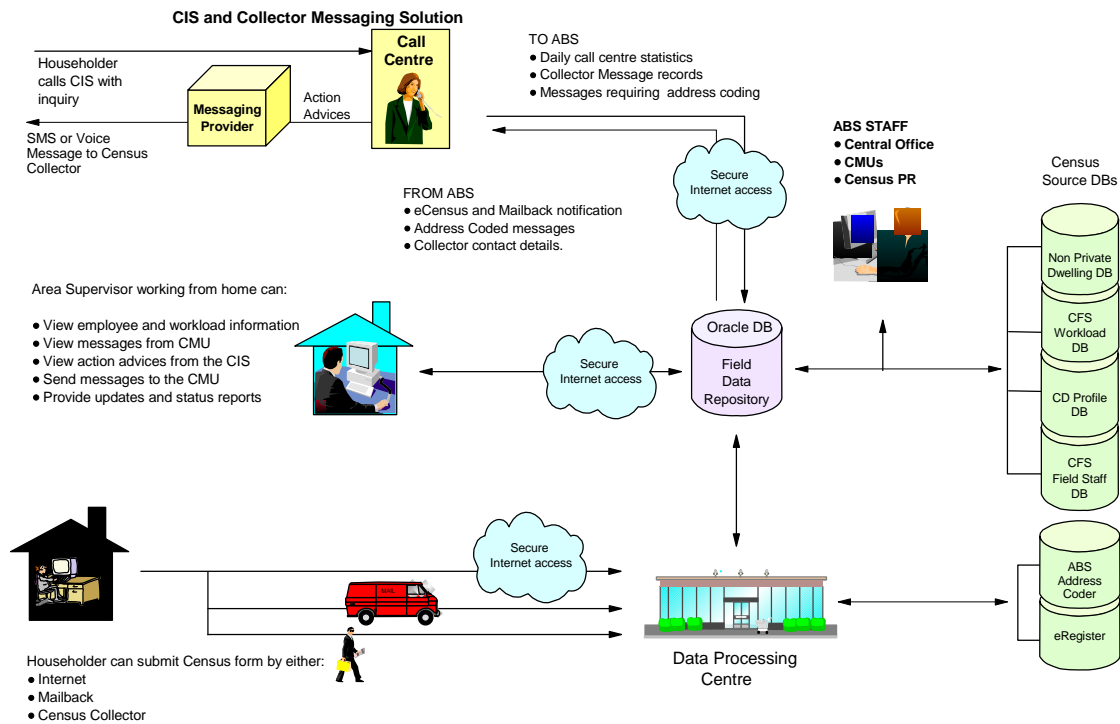
Outline of the Australian e Census

4. For the 2006 Census, Australia conducted a drop-off/pick-up census on a "de facto" basis, that is, for the purposes of the census count, people were counted where they were on census night not where they usually lived (usual residence data are derived from a usual residence question on the census form). A census enumerator delivered a census form to each address within a designated area and returned in the three weeks after census night to collect a form. The householder was encouraged to complete the census form on census night itself with reference to all people staying in the household that night.

5. For 2006 Census the householder was also given, in addition to the paper census form, a sealed envelope containing a unique 12 digit eCensus Number. If the person chose to use the Internet rather than return a paper form they could log on to the census web site and with a combination of the Census Form Number contained on the paper census form and the eCensus Number in order to gain access to the eCensus form.

13. The eCensus form was an online rather than a downloadable form. ABS examined the experience of other organizations in Australia and overseas that had used downloadable forms. The conclusion was, given the relative simplicity of the census form, an online form would be much more user friendly than a downloadable form.

14. Once the person has completed and submitted the eCensus form, a SMS (short messaging service) message is generated to the cell phone of the relevant census enumerator informing the enumerator that an eCensus form has been lodged and that the enumerator need not visit the house again.

15. This messaging system was integrated into an overall field communication system that also dealt with other census related inquiries from the public and the flow of management information between the census field staff and ABS (Australian Bureau of Statistics) management. An overview of this process is outlined in the diagram below:

CIS and Collector Messaging Solution

Householder calls CIS with inquiry

Messaging Provider

Action Advices

Call Centre

SMS or Voice Message to Census Collector

TO ABS
● Daily call centre statistics
● Collector Message records
● Messages requiring address coding

Secure Internet access

FROM ABS
● eCensus and Mailback notification
● Address Coded messages
● Collector contact details.

ABS STAFF
● **Central Office**
● **CMUs**
● **Census PR**

Census Source DBs

Area Supervisor working from home can:

● View employee and workload information
● View messages from CMU
● View action advices from the CIS
● Send messages to the CMU
● Provide updates and status reports

Secure Internet access

Oracle DB

Field Data Repository

Non Private Dwelling DB

CFS Workload DB

CD Profile DB

CFS Field Staff DB

Secure Internet access

Householder can submit Census form by either:
● Internet
● Mailback
● Census Collector

Data Processing Centre

ABS Address Coder

eRegister

16. Individuals within households and people in institutional dwellings were also able to complete an eCensus form even if other members of the household or institution completed a paper form.   Systems were also in place to monitor and handle duplicate forms from the one household (eg where there is both an eCensus form and a paper census form for the same household and/or people).

17. In addition to standard Internet security (128 bit SSL the same as used by banks etc), data transferred over the Internet were encrypted. As well, the ABS used a unique 12 digit eCensus number that acted as a password and was one of the measures the ABS used to prevent people logging on to the system and lodging spurious forms.

18. The basic business processes and systems described above were devised by ABS staff and tested in a number of small field test using basic infrastructure developed by the ABS in-house. This was to ensure that any solution could be integrated seamlessly into the overall census processes. .

19. For the census itself, the ABS outsourced major components of the overall solution. These included the development and deployment of the Internet application as well as the Internet infrastructure. As well, the Census Inquiry Service and elements of the field communication system were also outsourced. The major rationale behind outsourcing was the requirement to provide for a high peak load capacity on and around census night for the eCensus and the need to have ready access to the communications networks. ABS lacked both the infrastructure and the expertise in these aspects of the operation.

20. The diagram below outlines the elements of the eCensus undertaken by the ABS and those undertaken by IBM.
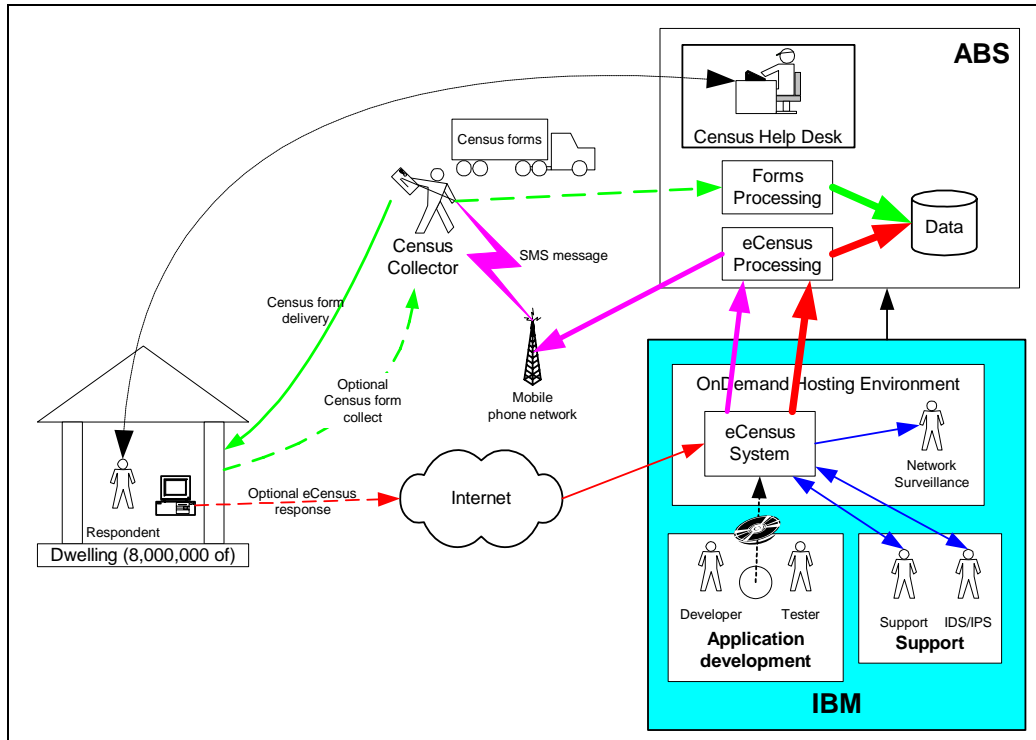
21. One additional benefit in the development of the eCensus and the associated field communication system was the opportunity to manage better the census field operations and contain the number of "non contact" dwellings. These are dwellings, which the enumerator has determined to be occupied, but for which no census form had been received. For the 2001 Census the proportion non-contact dwellings nearly doubled from the 1996 Census to around 2% of all dwellings for the 2001 Census. This is due to a number of lifestyle factors such as the increase in single person households and two person households with both partners working.

22. In previous censuses, the enumerator left a reminder note at each dwelling where a census form had not been collected requesting the householder to mail the census form back to the ABS and no further follow up was undertaken. There was no way of notifying the enumerator in a timely manner that a census form had been received from the dwelling.

23. For 2006 Census, enumerators were advised by SMS messaging to their cell phone whether a census form has been received either by Internet or by mail. Field managers were also able to organize and undertake intensive follow-up for those areas with large numbers of forms not received.

Performance of eCensus

24. The 2006 eCensus system was opened to the public just after 8pm on 27 July, with enumerators due to commence delivery of forms on 28 July. The first eCensus respondent submitted their online form at 20:29 on 27 July.

Take up rate

25. The number of Household forms submitted via the eCensus was **775,856**, representing take up by **9.0%** of the estimated dwellings in Australia. This take up is higher than the take up for the Dress Rehearsal. The final take up rate for the Dress Rehearsal was 7.9% of the 39,800 dwellings in the area.

1. eCensus Takeup
Household returns as a percent of estimated dwellings
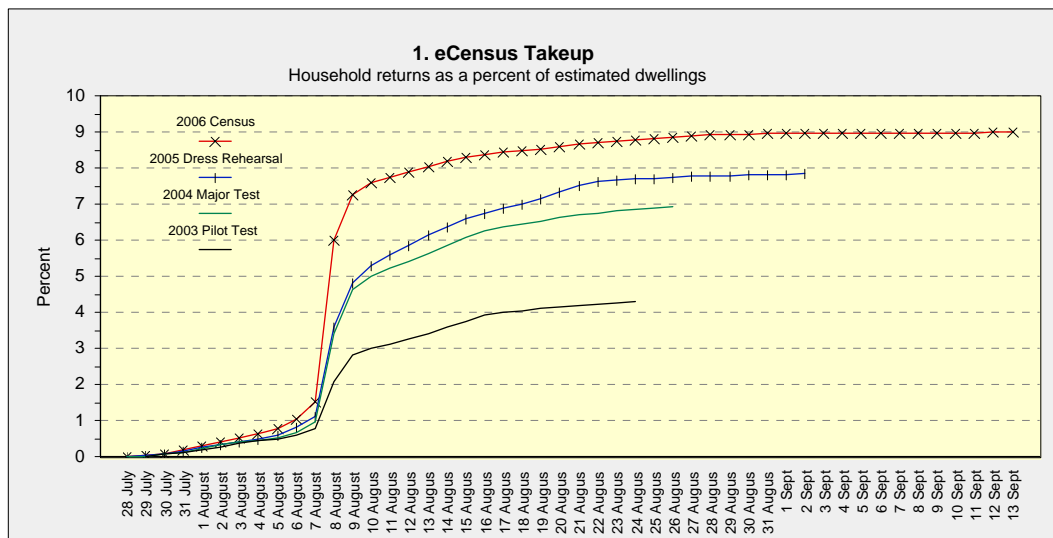
Take up by State/Territory

26. The following graph shows the take up in each State and Territory. As can be seen, take up is significantly higher in the Australian Capital Territory (ACT) than other states. Of the more populous States, Western Australia (WA) is slightly above the national average, with South Australia (SA), Tasmania (Tas) and Northern Territory NT below the national average. The most populous states, New South Wales (NSW), Victoria (Vic) and Queensland (Qld) are about the national average.

27. This take up largely reflects internet connectivity in each State and Territory which is highest in the ACT (82%), and lowest in Tasmania and the Northern Territory (44% each). The lower take up in the Northern Territory will also be partly due to the fact that the eCensus was not offered in discrete Indigenous communities, which represent a significant proportion of the Northern Territory's population.

2. eCensus takeup by State
as a percent of estimated dwellings
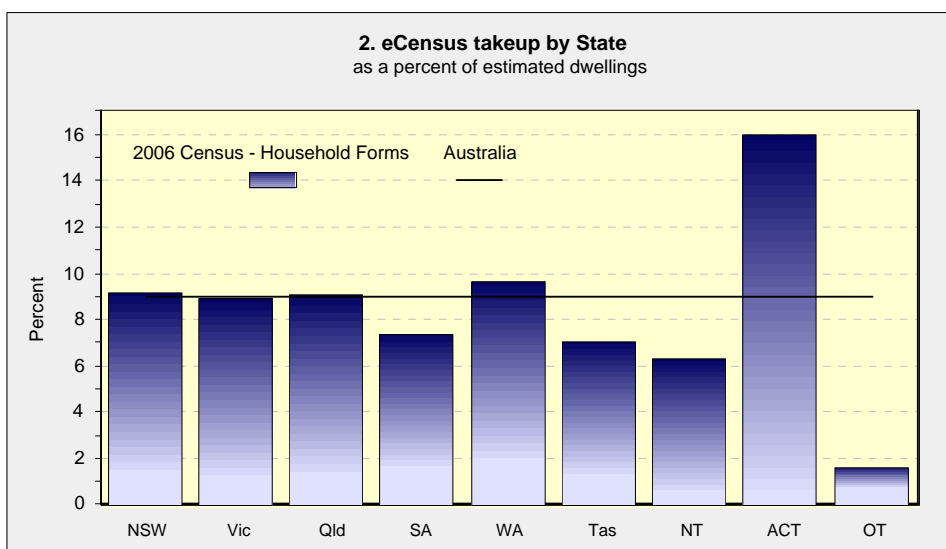
Submission pattern

28. The eCensus take up was characterised by a large spike in submissions on census night and a high level of use on the following day. Table 2 shows the distribution pattern for eCensus submissions across the different field phases, with comparison to the testing program. The most notable feature of this table is the proportion of forms submitted on census night, which at 40.5% was significantly higher than was seen in any of the field Tests (22-26%). There was also a lower proportion of submissions on the following day, and throughout the collection period.

29. The different pattern seen in the census may be due to the awareness raising public relations campaign, which emphasises census night. It is also possible that the higher proportion of submissions seen in the collection phase of the Tests was affected by the type of staff employed as enumerators: it is possible that enumerators employed in Tests are of higher quality (due to the smaller number of positions to be filled) and are therefore more diligent in collection and follow-up.

Table 2. eCensus Submissions, by field phase				
Phase	**2006 Census**	2005 Dress Rehearsal	2004 Major Test	2003 Pilot Test
Delivery period	16.9	14.6	14.0	18.6
Census Day, prior to 6pm	9.3	6.9	9.4	7.2
Census Day, after 6pm	40.4	24.6	25.8	22.6
Day after Census Day	14.0	15.4	17.7	17.2
Collection period	19.3	38.6	33.2	34.4
Total	100.0	100.0	100.0	100.0
Estimated dwellings	8,617,000	39,800	20,600	6,450

30. The following graph shows the daily distribution of eCensus submissions. Comparative data from the Dress Rehearsal are also shown. Apart from the expected census night spike and strong activity in the 'shoulder' period (especially the following day), this graph also shows that a small proportion of people complete their forms on the Tuesday prior to census night (1 August), as well the last weekend on which enumerators were active in the field.

3. eCensus Returns
Daily takeup as a percent of total returns
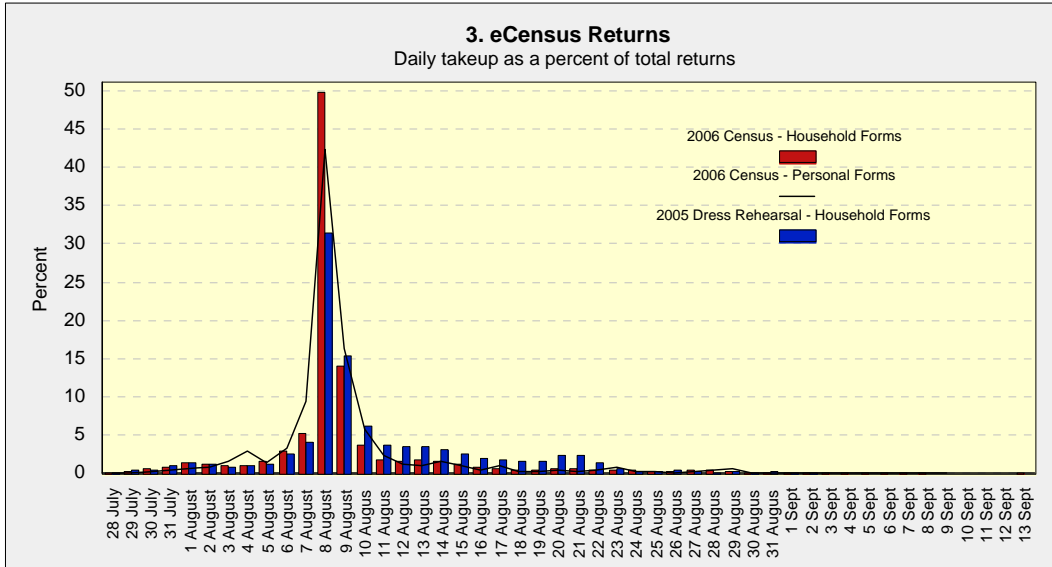
31. As noted previously, the time of day data shows a large peak in submissions in the evening of census night. The graph below shows that during the delivery period, there is increasing use throughout the day, leading to a peak in the early evening. census day itself shows consistent use throughout the day, with a large peak starting to build from 4pm onwards, tapering off after 11pm.
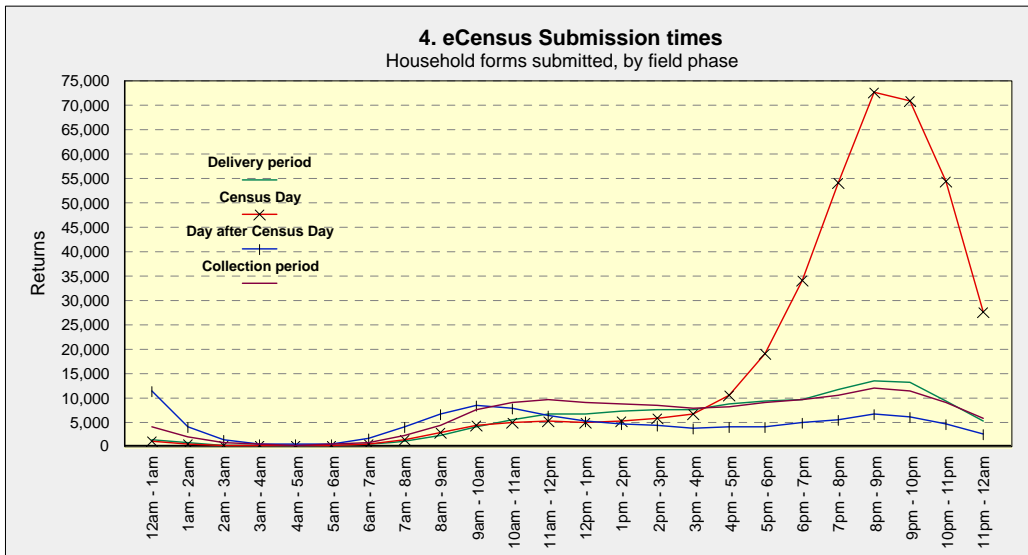
4. eCensus Submission times
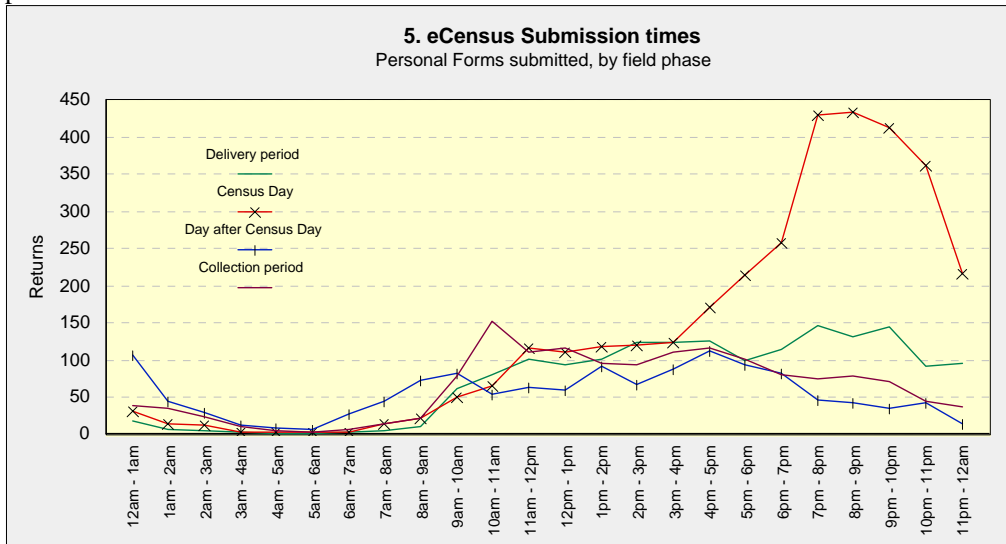Household forms submitted, by field phase

Note that the green 'Delivery period' and purple 'Collection period' lines represent an aggregate period, while the red 'Census Day' and blue 'Day after Census Day' lines each represent a single day's activity

Personal Forms

32. The eCensus equivalent of a Personal Census Form was available to people in Non-Private Dwellings such as hotels, boarding houses, etc. A total of **7,838** Personal Forms were submitted via the eCensus (see graph 3 above for distibution).

33. The following graph shows the time of day for submission of Personal Forms. The pattern is similar to that for household forms.

5. eCensus Submission times
Personal Forms submitted, by field phase

Delivery period
Census Day
Day after Census Day
Collection period

Returns

450
400
350
300
250
200
150
100
50
0

12am - 1am
1am - 2am
2am - 3am
3am - 4am
4am - 5am
5am - 6am
6am - 7am
7am - 8am
8am - 9am
9am - 10am
10am - 11am
11am - 12pm
12pm - 1pm
1pm - 2pm
2pm - 3pm
3pm - 4pm
4pm - 5pm
5pm - 6pm
6pm - 7pm
7pm - 8pm
8pm - 9pm
9pm - 10pm
10pm - 11pm
11pm - 12am

Note that the green 'Delivery period' and purple 'Collection period' lines represent an aggregate period, while the red 'Census Day' and blue 'Day after Census Day' lines each represent a single day's activity

Critical Success Factors

34. Some of the critical success factors are listed below:

1. Capacity of final solution.

35. The final eCensus solution used for the 2006 Census was sized to deal with up to 25% of the population submitting their census form online over the census period and with 20% doing so on census night. Extensive volume testing and systems tuning was undertaken prior to the census. Volume testing was also undertaken by an independent consultant to verify the figures being supplied by IBM. Volume testing covered both the ability of the systems and infrastructure to handle both sustained load as well as spikes.

36. During census operations, the eCensus system was closely monitored for any unexpected surge in usage. The capacity of the system was never really put to the test – with peak load on census night reaching only 15% of capacity.

37. There were contingency plans in case the capacity was exceeded. The first of these prevented any further people logging on to the system once a certain load was reached. This was set at a level that ensured that people already logged on could continue to complete their forms without any significant delay. Those people denied access would receive a message that the system was unavailable and to log on later. Public relations messages were also ready to inform the public of any capacity constraint on the eCensus site and that they may need to postpone logging on. These contingency plans were never required to be implemented.

2. Perceived and actual security of the eCensus system and the internet

38. The eCensus security solution was compliant with the Australian Government Protective Security Manual (PSM) and government IT security guidelines contained in ASCI 33. It utilised proven security technologies such as multiple levels of firewalls, intrusion detection

and protection devices and application servers built on hardened operating systems. The solution made extensive use of strong encryption to ensure the confidentiality of respondent data. Systems management policies were put in place to ensure that IBM employees had only tightly controlled access to eCensus infrastructure components.

39. In order to ensure that IBM staff were not able to access eCensus data all submitted and unsubmitted data were encrypted prior to being saved. The data were encrypted using a dedicated hardware device initialised with a private key by ABS staff. IBM did not have access to the private key. The encryption used is rated as secure until 2015 (that is using conceivable processing resources it would take at least until 2015 to break the encryption). Application code reviews and technical auditing were in place to ensure that IBM did not copy eCensus data in any form. At the conclusion of eCensus ABS security staff oversaw the secure deletion of eCensus data from all IBM systems. This deletion was carried out in accordance with Australian Government standards for the deletion of security classified data. Additional encryption techniques were used to protect the data while in transit over the internet or within IBM's network.

40. ABS IT Security section and the Australian Department of Defence were involved in the solution design. It was reviewed at numerous stages by an external security consultant. Prior to the system going live the Australian Department of Defence were engaged to conduct vulnerability assessments of the solution. External security consultants conducted detailed application code reviews and carried out active penetration testing.

3. The capacity of Australian ISPs (Internet service providers) to handle increased levels of Internet traffic on and around Census night.

41. An inability of Australian ISPs to handle high levels of Internet traffic generated by use of the eCensus could have severely impacted upon the take up rate of the eCensus. Should the services of a major ISP have been unavailable, a large number of people would have been unable to connect to the internet, and therefore unable to log on to the eCensus form. This may have resulted in respondents choosing to revert to a paper form, or not completing a census form at all.

42. A strategy for the notification of ISP's in the lead up to the 2006 Census was implemented, in an attempt to manage some of the issues arising from increased internet traffic generated by the eCensus.   There was a problem at the start of the census operations with some smaller ISPs not including the eCensus address in their indexes making it impossible for their customers to access the eCensus site.

4. Malicious attacks on the e Census web site

43. A denial of service attack would not only reduce the ability of people to lodge their census form on line but could also have had an adverse affect on the image of the ABS and of the census. ABS along with IBM implemented a range of measures to contain any such attacks. These include continuous monitoring to detect possible attacks with the options of locking out source ISPs for up to thirty minutes or provision of additional access routes to the census site.   There were a few occasions where unusual levels of activity were detected from certain IP addresses that lead to these addresses being denied access.

44. There was also a contingency that, in cases where these attacks could not have been dealt with quickly, public relations messages would have firstly assured the public that their census information is secure and secondly provide information about alternatives such as delaying using the e Census system or using the paper census form. This contingency was not required.

5. Integration of eCensus into other census systems and procedures

44. The successful implementation of the e Census was reliant on the successful integration of the eCensus system into existing census systems and procedures. This was to both maintain the quality of the census and the cooperation and support of the public.

45. For census field systems, this included the successful operation of the Census Inquiry Service (that handled inquiries from the public about the census) and the census field communication systems, to ensure that every household and person was counted and that households that completed an eCensus form were not followed-up unnecessarily.

46. Changes were made to the census processing system to accept eCensus forms and integrate them into the processing streams. Extra checks were implemented for duplicated household and person records and a method has been implemented to determine which of the duplicated records was retained.

6. Ease of access and usability of the eCensus system

47. A key requirement for a successful eCensus was that the system should be readily accessible, easy to access and easy to use. All households were given an eCensus number that would allow them to access the eCensus system. The eCensus system was compatible with nearly all browsers – both commercial and open source and could be used via by both dial up and broadband connections. A great deal of attention was given to ensuring that the form itself was easy to use and extensive usability testing was undertaken. As well, a form that was compatible with the major screen readers used by the blind was also made available. In addition to help facilities built into the form itself, the ABS also provided a technical telephone help line though the Census Inquiry Service.

48. There was an opportunity on the eCensus form for users to provide comments. These comments were overwhelmingly positive, in particular about how easy the form was to use. There were a small number of complaints from people who had ongoing difficulties with the eCensus system and it was not always possible to identify the source of the problem.