Census 2006: capitalising on IT

John Paynter

Datamail

jpaynter@datamail.co.nz

Gabrielle Peko

Department of Information Systems and Operations Management
University of Auckland

g.peko@auckland.ac.nz

Abstract

The 2006 census in New Zealand introduced a number of new initiatives in order to improve the processes and as steps towards a future involving e-government. An online census form was made available as an alternative to paper-based forms. In order to control this and other census processes an automated system of text messaging was used to send information to field staff about census form collection and delivery. They were also advised when forms had been submitted electronically or via the post in order to avoid making unnecessary trips to pick them up. The uptake of the e-forms was not as high as expected. There were problems with the text messaging system. The lessons learnt are reviewed in order to improve the processes for future censuses and other services provided by the government.

Keywords: Census, text messaging, online forms, e-government

1 Introduction

A census is an official count. It can be contrasted with sampling in which information is only obtained from a subset of a population. As such it is a method used for accumulating statistical data, and it is also vital to democracy (voting). Census data is also commonly used for research, business marketing, and planning purposes. In New Zealand a census is held every five years. It is a snapshot on the chosen day when the number of people and dwellings (houses, flats, apartments) counted. Everyone in the country on that day is asked to complete census forms. There are two census forms. The blue Individual Form (IF) must be completed by everyone in your household on census day. The brown Dwelling Form (DF) must be completed by one person in your household. For the 2006 census an option was introduced to complete the forms on the Internet. Other initiatives included sending text messages about this process, amongst other things to the enumerators (collectors) whose job it is to collate the information in the field.

Information technology, especially the Internet, opens possibilities of using methods to distribute information and deliver services on a much grander scale (Paynter and Fung, 2006). It can deliver government services and encourage greater democracy and engagement from citizens. Governments around the world are exploring the use of web-based information technology (Grönlund, 2002).

Since the mid-1990s governments have been tapping the potential of the Internet to improve governance and service provision. "In 2001, it was estimated that globally there were well over 50,000 official government web sites with more coming on-line daily. In 1996 less than 50 official government homepages could be found on the world-wide-web." (Ronaghan, 2002).

Along with the rapid growth of technological developments, people demand high quality services that reflect their lifestyles and are accessible after normal office hours from home or work. Thus, the goals of delivering electronic government services are to simplify procedures and documentation; eliminate interactions that fail to yield outcomes; extend contact opportunities (i.e., access) beyond office hours and improve relationships with the public (Grönlund, 2002).

This paper looks at how technology has been progressively introduced to recent censuses to improve the information flow to and from field staff through the use of automated text messaging. It also examines the uptake of electronic submission of census forms. The lessons learnt from the 2006 census are examined in the light of showing how government can further capitalise on IT and the Internet in particular to promote democracy services.

2 Background

Census-taking began in China and the Middle East. One of the earliest recorded censuses took place in the Babylonian Empire nearly 6,000 years ago. Early censuses are mentioned widely in early Middle Eastern literature, with references to them in a number of places in the Bible.

Censuses of population were first taken in England and Scotland in March 1801, Ireland in 1811 and Australia in 1828. In the USA the census is undertaken every 10 years. The US Census 2000 project spanned 13 years at a cost of $65 billion (Gido and Clements, 2006 p147). It is largely based on mailbacks with census employees personally visiting non-respondents. The first New

Zealand census was undertaken in 1851, although this census excluded Māori (Statistics New Zealand, 2006a).

In New Zealand several acts of parliament have formed the legal basis for the collection of statistical data and census taking that has developed over the years. The most recent of which is the Statistics Act 1975. It clarified that the information contained in returns is to be used for statistical purposes only. It also specified which particulars it is mandatory to collect in the census and which particulars are able to be collected if the Government Statistician considers it in the public interest to do so. It also guaranteed the census to be free of government influence.

2.1 The Use of Technology

The 1921 Census marked the first occasion on which automatic sorting and counting machines were employed in New Zealand, enabling the major portion of census compilation to be carried out mechanically. The system installed for this census was purchased from the United States, which had been employing mechanical tabulation for census work since 1870.

For the 1966 Census, sorting machines were replaced by computers. Statistical tables were also produced by computer for the first time and results became available much earlier with a large number of additional cross-classifications of the census data being possible. The use of punchcards for each individual and dwelling was continued until 1976 when an automatic, electronically-based system was introduced. Mechanical tabulation has been replaced by electronic data capture and handling as the speed and capacity of computing technology has improved.

In 1996 the scanning and imaging of census forms was introduced, further demonstrating that Statistics New Zealand was now fully immersed in the era of information technology, with analytical tools and information at a level incomprehensible to the department of earlier years.

2.2 Enumeration

In the 2001 census though the process of distributing and collecting forms (enumeration) had hardly been changed. Enumerators within each district would hand deliver forms to each household (one dwelling form and one individual form for each person expected to be present on census night). Each form would be coded with an identifier made up of District, Subdistrict, Meshblock and Dwelling (the individual forms had Person ID added on collection). This ID was recorded in a field book along with any comments including the address and best pick up time. After Census night the households would be visited again to collect the completed forms. Up to three visits would be made in each of the delivery and collection phases. On the third unsuccessful visit prior to census night a default number of forms (one dwelling, three individual) would be left. After census night an envelope would be left on the third unsuccessful collection visit. At the end of the enumeration phase the District Supervisors would send follow up letters and / or visit non-respondents. Once District offices were closed

five weeks after the census, the central census office would follow up non-respondents.

3 Census 2006

Census 2006 would see the introduction of two technological innovations. The first was the adoption of Internet-based census forms (Statistics NZ, 2004). This would enable the dwelling and individual forms to be submitted electronically via the Internet by the householder and individuals. The second, in part necessitated by the first, was to automate the flow of information about the forms submitted either electronically or via post to the enumerators (collectors). On the basis of the Internet ID on the census form the enumerators were texted via webmail to their census cellphones with the details of forms submitted.

3.1 On Line Census forms

On delivery of the census forms to the household the enumerator would ask the "Hi-Five" questions. These included whether or not members of the household would like to submit their forms on-line. If one of more individuals indicated that they might want to do this, then they were given an Internet PIN for everyone in the household to use. The household ID (District, Subdistrict, Meshblock and Dwelling) forms the Internet ID for the entire household. A sealed slip was given out to the household. This contains the PIN. Although everyone within the household used this one combination of ID and password, their information would still be secure as each individual's had to be entered within the one session. That is, a session could not be saved and resumed. However the different individuals could enter their information at different times. Once each dwelling or individual form was completed the system would batch the submissions and text the enumerator on a thrice-daily schedule. They would get a cumulative report of the number of forms of each type submitted for each household.

3.2 Text Messaging

The computer Field management system (FMS) records the contract and contact details of each of the enumerators. With the exception of some of the high-density apartment buildings in the CBD, each subdistrict is assigned a single enumerator. The census cellphone number of each enumerator is recorded whether it be their own phone (for which an allowance is paid) or one provided for the census. During the census period FMS can text messages to the enumerators. This is done on the basis of the Internet ID (District + Subdistrict) where the message is sent in reference to a pre-coded form or the address when initiated by an individual who perhaps does not have a form.

These ACTION messages comprise one of four types. An instance of each type is shown below.

```
0010498,6320201001W,3
Bombay,NOTIFY,(I:1e),dispatched more forms
```

```
0010499,6320201,5         Bombay,ACTION,,housesitters
staying here; deliver forms
0010503,6320201006F,11  Bomba,INFO,,Destiny  Child
will need assistance.
6320201004P,9 Bombay,OFFICE,(D:1,I:3,A:1),
```

The first part is the unique message number; this is followed by the Internet ID (632=District, 02 = Subdistrict, 01 = Meshblock, 001 = Dwelling, W = Checkletter; this is followed by the address; the message type (NOTIFY, ACTION, INFO, OFFICE); the type (D = Dwelling, I = Individual) and number of forms; and lastly any textual information. The forms may be either in English (e) or Maori (m) – the two official languages in New Zealand.

The enumerators were required to enter any messages in their field book against the line for that particular dwelling. In the case of ACTION messages they were to deliver the required number of forms to the dwelling. OFFICE messages denote that the forms have been received via the Internet (online submission) or have been mailed. The enumerator would update the forms received in the OFFICE column of the field book and could check this against those delivered to see if any further forms needed collection for that address.

These messages were also available in the action log. The district supervisor could print these to check against the enumerator's field book when they made field checks and when the enumerator brought the field book in with the boxes containing the forms they had picked up at the end of the field phase of the census.

4   Results

Trials conducted a year before suggested a 20% uptake of the electronic census form option (Statistics New Zealand, 2005.

Field staff were instructed not to oversell this option as there were fears that 30% uptake would be too much for the computer system. During the delivery phase staff reported a high uptake in some areas (as much as 80%, particularly in inner city suburbs and the CBD). Although the forms were to be completed as of census night (March 7th), the system was available from 20 February (the beginning of the delivery phase) till the end of March. The peak period though was on census night and few problemsi were encountered with the use of the Internet forms. However some of the OFFICE messages to the enumerators were not received during this time. It should be emphasised that the forms themselves were received in the census system. There were also performance and other problems with the Field Management System (FMS), especially considering that access to it from the district offices scattered throughout the country was via dial-up lines (Jackson, 2006). Exacerbating this, many more mailback forms were received at the central office in Christchurch and the district offices. Each of these was to be entered in the system so that FMS could text the enumerator with the corresponding OFFICE message. This was not feasible in the District offices and the information entered at the central office lacked addresses (omitted to cut down the time it took to perform data

entry), so it was difficult to determine the correct dwelling when the forms had been miscoded.

At the end of the field phase only 8% of the completed forms had been received via the Internet option. The author questioned people during the collection phase of the census (when doing doorstep checks). Many of those who had requested PINs so that they could do the census online stated that as they had the paper-based form (given out both to record the Internet ID, and as a back up in case the electronic version failed) they found it easy to complete the paper form, particularly with family members sitting around the table. Conversely in the large flatting situations typified in some suburbs, only having a single PIN made it difficult for those who would elect to complete the forms at work and other places where they had Internet access.

Another situation again arose in the case of Non-Private Dwellings (NPDs) such as hostel, hospitals and hotels where each individual was given a separate PIN. However the uptake in such places tended to be low. Of note also, it was not intended to text the enumerators information on the forms submitted electronically by individuals in the NPD. The reasoning for this would appear to be that the message would not be received in time before the enumerator visited locations such as hotels and motels the morning after the census.

In contrast over 10% of census forms had been mailed in. These and the ones collected by the enumerators and district supervisors had to be dispatched to Christchurch to be scanned. Some preliminary results for the census (e.g. the overall population count, based on analysis of the field books) were available at the end of May (Statistics, 2006b), but it would take three months for the forms to be scanned and the data were to be made available on December 6.

The use of the field books too was error prone. Enumerators could wander out of a meshblock and even into another subdistrict or district. Such recording errors would have to be corrected in later phases of the census. This would be very hard to do if it did not occur at the district level. For instance if the online option was taken there was no mechanism for the District Supervisor to correct the return – it could only be logged as an enumeration event.

The enumerators texted, emailed or phoned their daily delivery and collection statistics to the district supervisors. Although these were to be entered in FMS problems arose and a largely manual system was used. This resulted in further time delays in receiving information. Thus it would be hard say to realise more staff were needed when delivery and collection were encountering problems.

5   Discussion

5.1   E-census

Lessons were learnt that will improve the participation in the next census. The risks of the online version were recognised in the trial held a year earlier. Emphasis was

made on training of the field supervisors and the collectors. The collectors were given access to the online census facility, to enable them to become familiar with the online option before it went live. This was invaluable as an aid to their understanding of how the option would be viewed by respondents prior to standing at their doors and offering that option (Statistics, 2006) Additional training was also introduced for the field collectors to ensure they understood what is required of the respondent, regardless of which option is chosen. This was to ensure, firstly, that the online option is communicated correctly to the respondent on the doorstep; and, secondly, that the field communication systems integrating the online option with the paper option were correctly performed. The training theme continued with field supervisors, who need to be able to use the field operations monitoring systems. These are the mechanisms used to monitor how the team in the field responds to the text messages they receive about the lodgement of Internet forms and related helpline actions. The complexity of these systems and the need to integrate with existing systems for the paper form collection process posed a real challenge to ensure seamless operations. Unfortunately the failure of FMS (Jackson, 2006) meant that this integration was not a success. The field team became distrustful of the accuracy of the information provided about forms collected by the online and mailback options. This made their jobs harder in the collection phase and then when the collectors returned the forms. Information about mailbacks, privacy envelopes, e-forms and hand-collected forms had to be combined. Subsequent to this follow-up letters were sent to apparent non-respondents but again the possibility existed that the forms had been received by mailback or electronically but that this was not communicated to field staff

5.2   E-government

It has been claimed that the low penetration of broadband technology in New Zealand is a limitation to the spread of electronic services. However news of the proposed unbundling of the Telecom local loop monopoly was released prior to the 2006 Budget (Budget, 2006). Clearly the infrastructure is in place for the participation of New Zealand citizens in e-government. However the uptake by both the citizens, as shown in the census, and government agencies is low. For instance, only one council, Auckland Regional Council, provides an on-line forum for discussion and sharing of ideas (Paynter and Fung, 2006). None of the local government sites provide any e-democracy although some sites have put up information about Elections 2004. Dunayev (2005) used an automated tool to analyse all the local government web sites. He concluded that the sites did not appear to have matured sufficiently to meet the goal of online local government elections in the next cycle (2007). Some of the obstacles to e-voting, such as trust, are outlined in Sharkey and Paynter (2003) and steps towards an e-voting transition in Paynter and Peko (2005). This included the use of e-services in the census and in the local body elections – both less potentially risky and lower profile than a general election.

Other countries that have an e-census option include Switzerland where 4.2% of the population took part in the E-Census. 11% who began to fill out the form online interrupted the process before reaching the end. The E-Census homepage received 238,000 visits. Of those households that took a look at the Web site, only half actually filled out and submitted their questionnaires online. This shows the wide gap between simply searching for information over the Internet and carrying out a complex transaction online (Swiss Statistics).

5.3   Preparing for the Future: Capitalising on IT

The information and services provided on-line via by governments is constantly undergoing evolution partly driven by innovations in Information Technology (IT), partly by government wishing to leverage this tool and to a minor extent by user demand. Although the uptake by citizens to undertake the census online option was poor, this can be improved as the Internet becomes more pervasive and lessons learnt from the 2006 census incorporated into the training and testing in 2011. Given the problems in contacting households, in particular in high rise apartments and walled communities, the Internet offers a viable option.

Use of the field books is cumbersome and error prone. There are also delays in communicating such information captured on hard copy paper. It would be advantageous to use an electronic notepad allied with a Global Positioning System (GPS) to minimise recording errors (such as failing to give the complete internet ID or moving out of the correct meshblock) and maximise the responsiveness to the delivery and collection statistics recorded. Surveylab (www.surveylab.co.nz) is an example of such a system used for GIS information (Stuff, 2006).

6   Conclusion

For the most part, although the technical side of the electronic census worked, its uptake by citizens was low. This would suggest that further moves towards e-democracy in the form of electronic voting would be premature in terms of public acceptance and uptake.

7   References

Australian Bureau of Statistics (2006) Welcome to eCensus. Accessed on 26 September 2006 from http://www.census.abs.gov.au/eCensusWeb/

Budget (2006) Government moves fast to improve Broadband. Accessed on 3 May 2006 from http://www.beehive.govt.nz/ViewDocument.aspx?DocumentID=25636

Clements, J. & Gido, J. (2006) Successful project Management, Thomson, Mason, OH (3rd Ed), p147.

Dunayev, A (2005) Electronic Local Government Elections in New Zealand. Unpublished BCom (Hons) dissertation, The University of Auckland, New Zealand.

Grönlund, A. (2002). Electronic Government: Design, Applications & Management. Hershey: Idea Group Publishing.

Jackson, R. (2006) Census system overload but no data lost, Computerworld, 20 March, 2006. Accessed on 3 May 2006 from http://computerworld.co.nz/news.nsf/UNID/0A98536DAE494843CC257133007AEE8B?OpenDocument&Highlight=2,census

Paynter, J & Fung, M. (2006) 'E-Service Provision by New Zealand Local Government', In:Ari-Veikko Anttiroiko, & Matti Malkia, (ed.), *Encyclopedia of Digital Government*., Hersey, Idea Group Publishing, 718-725, 2006

Paynter, J & Peko, G. (2005) e-lections and the price of democracy, Proceedings of the 40th Operations Research Conference of New Zealand, Wellington Dec 1-3, 2005, 145-154

Sharkey, E. & Paynter, J. (2003) Factors influencing the uptake of Online voting in NZ. CHINZ '03: the 4th annual conference of the ACM Special Interest Group on Computer-Human Interaction, New Zealand Chapter Dunedin. 3-4 July 2003, 121-122

Statistics New Zealand (2004) On-line Option for 2006. Census Accessed on 22 June 2006 from http://www2.stats.govt.nz/domino/external/pasfull/pasfull.nsf/web/Media+Release+2006+Census:+On-line+forms+October+2004?open

Statistics New Zealand (2005) Using Both Internet and Field Collection Methods for the 2006 Census of Population and Dwellings Accessed on 19 September 2006 from http://www.stats.govt.nz/census/2006-census/methodology-papers/internet-and-field-collections.htm

Statistics New Zealand (2006a) Introduction to the Census (2001) - reference report. Accessed on 20 June 2006 from http://www.stats.govt.nz/census/2001-census-technical-info/2001-introduction/default.htm

Statistics New Zealand (2006b) 2006 Census of Population and Dwellings – Provisional Counts. Accessed on 29 May 2006 from http://www.stats.govt.nz/products-and-services/hot-off-the-press/2006-census/2006-census-provisional-counts-2006-hotp.htm

Surveylab (2006) Customizing applications for ike, Accessed on 27 September 2006 from http://www.surveylab.co.nz/~downloads/CUSTOMIZING%20IKE.pdf

Swiss Statistics (2001) Experience with the E-Census Accessed on 26 September 2006 from

http://www.bfs.admin.ch/bfs/portal/en/index/themen/volkszaehlung/uebersicht/blank/zur_erhebung0/erfahrungen_mit_e-census.html

[i] If pop ups were disabled the forms could not be completed as there was a pop up confirmation message at the end of the submission stage. Some collectors had either not coded the forms at all or had omitted the check letter. Where there were more dwellings in a meshblock than expected, although an extra 20% loading was catered for, an overflow book had to be used and there was no corresponding Internet ID. i.e. all the check letters were coded as 'X'. In these circumstances the respondents could not use the Internet option for completing the forms.