The Expert Group Meeting on Effective Use of IT in Population Censuses

Bangkok, Thailand on 10-12 December, 2007

By

Mr. Boonsom   Suwannakhot

Ms. Angkana Hatthakitkosol

Ms. Pannee Pattanapradit

National Statistical Office

Bangkok, Thailand

Use of IT in Thailand 2000 Population and Housing Census

I. Introduction

National Statistical Office Thailand is the main governmental agency entrusted with responsibility for Thailand's statistical activities including the collection and compilation, dissemination of fundamental statistics, providing recommendations on statistical related matters, organizing training courses in statistical methods, computer data processing as well as the statistical data bank of the country, co-coordinating with international organizations in terms of statistical techniques and operating any other activities which are commanded by the Prime Minister of the Cabinet.

II. The 2000 Population and Housing Census

National Statistical Office has had long experience in undertaking national population and housing censuses, apart from its responsibility of collecting, interpreting, and disseminating statistics in other fields.

The 2000 Population and Housing Census is the tenth population census and the fourth housing census of Thailand. The National Statistical Office (NSO) carries out the Population and Housing Census every ten years in accordance with the international standard to collect basic data on the number and characteristics of population and housing throughout the country. The data are used for planning and setting policies of the government and private sectors. The census data are also used for population projections.

Methodology

The data collection was interviewing method. All provinces except Bangkok, school teachers were used for supervisors and interviewers in both municipal and non-municipal areas while temporary employees in Bangkok were hired for both supervisors and interviewers. There were 40,000 enumeration districts. The total of 6,000 supervisors and 40,000enumerators were used for the field enumeration.

The sample census enumeration technique was used to reduce cost. All persons and households in every province were listed and simultaneously enumerated with the short form

questionnaires except for the sample households (20%) which were enumerated with the long form questionnaires in every province including Bangkok.

The questionnaires

There were 15 questions in the short form and 35 questions in the long form.

Data processing and ICR technique

The 2000 Census had adopted the ICR as the core technology in processing the detailed data in order to improve timeliness of data reporting. This technology has proved to be reliable and helped reduce data processing time. In doing so the questionnaire forms (16 million forms) need to be captured and translate into ASCII files as soon as possible for further edit, tabulation and publication. The NSO decided to use the new technology, the Intelligence Characteristic Recognition (ICR). The scanner of ICR can read the questionnaires at very high speed with the program that can translate the number and alphabet of hand writing into the ASCII file for further processing rather than data keying system**.** However, there have been some problems. For example, the distribution and return of census questionnaires must be done with caution, especially from the remote areas. Some questionnaires were found to be wet of rain or wrinkled and some were recorded with bad hand writing, and these had affected the quality of ICR.

Data dissemination

The 2000 census has been carried out under conditions of financial stringency (Asian Financial Crisis). Data sets were made available as follows:
- Publications (Preliminary report, advance report, the Final reports of all provinces)
- On-line data dissemination, including summary statistical information on the Website:
   http://www.nso.go.th
- CD-ROM, diskettes

III. The 2010 Population and Housing Census (Preparation and Planning)

Digital Mapping

Mapping work for the 2010 Census will begin as early as 2008 and be completed in early 2010. The main objective of mapping is to provide the latest information on the census frame to enable enumerators to carry out their task in a timely and precisely manner. The

operational work involved in relation to updating the maps was fully carried out by the Provincial Statistical Offices (PSO). The MapInfo was used for digital mapping.

Data processing

For data processing stage of the next Population and Housing Census, the NSO will be able to build a strong ICT infrastructure to enhance its system of data processing and dissemination. Image scanning system is a technology for efficiently managing the data capture for large-scale surveys and censuses. It provides accurate, timely and reliable data capture from the census forms with less human intervention than needed for manual data entry. Because the census requires large-scale data collection, the image scanning system which needs to be installed must be sufficiently efficient to process the data within the required period. For such a large investment, the possibilities for using the technology after the census is completed will have to be considered.

By 2010, The NSO is currently planning to use the Intelligent Character Recognition (ICR) for data capture and after series of testing it will have to be considered the system.

The NSO decided to use the Intelligence Characteristic Recognition (ICR). The scanner of ICR can read the questionnaires at very high speed with the program that can translate the number and alphabet of hand writing into the ASCII file for further processing rather than data keying system.

The reasons for using ICR are as follow :

1.   The ICR technology that the scanner can read the questionnaire forms at a very high speed and with the program that can translate the number and alphabet of hand writing into the ASCII file for further processing by mainframe computer.

2.   Accuracy, the ICR can read the information more accurate than data keying.

3.   Reduce cost and space

4.   After processing the census data, the ICR can improve the timeliness in processing other the NSO's census and survey questionnaires.

Dissemination through the Internet

Internet is widely emerging technology that is gaining a strong popularity. The Internet plays an important role in the growing information technology. In statistics, the Internet has shown potential advantages in both statistical production system and a dissemination system.

Using the Internet in the NSO is viable through a long term plan. As an internet user currently, the technology has shown promising benefits to statistical system. As an Internet provider, the technology will contribute to the NSO in a way that statistical information become more accessible.

The NSO plans to improve the data dissemination system, and develop on the Internet important various services are accessible by users as follows:

- Website: http://www.nso.go.th

 This might be an alternative for publishing and presentation works for statistics.

- Data Warehouse

The NSO plans to improve the data dissemination system. The result of this improvement will help data users to meet the need and to access data and information more conveniently through the on line.
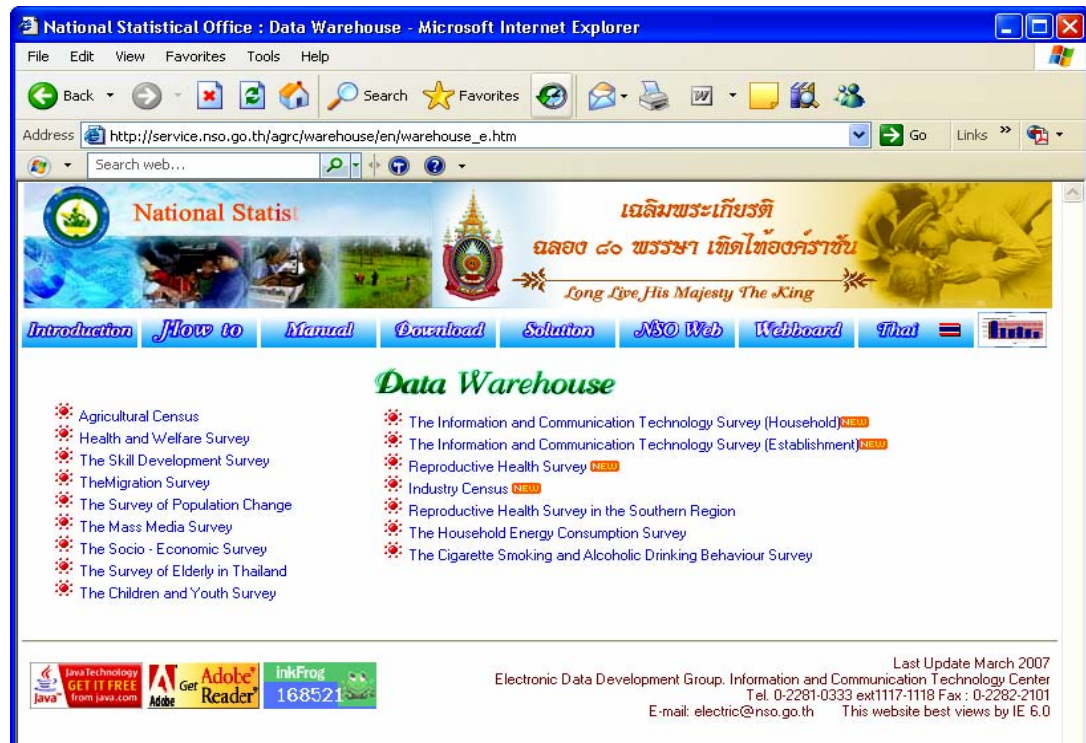
- Web Service (to/be consider)
- Geographical information system (GIS)

Conclusion

Census management strategies emphasize early planning, testing, and continuous evaluation of census operations to reduce the periodic census burden and improve quality.

Since Thailand has had experience in using the ICR technology for the previous census, it will be used again in the 2010 Population and Housing Census and more ICR system will be set up in regional centers. The NSO will be able to build a strong training of trainers, implementing a public awareness and educating campaign on the 2010 Census. The lessons learnt from previous census, together with good planning and preparing will make the 2010 Population and Housing Census to be more successfully In planning dissemination, setting up and design a website for census data (results) is one of dissemination, system.