

Exploiting New Information Technology in Indonesia Census Operations

Sugito Suwito

**Central Bureau of Statistics (BPS)
The Republic of Indonesia**

**18th Population Census Conference
Program on Population, East-West Center
Honolulu, Hawaii USA
26-29 August 1998**

Exploiting New Information Technology in Conducting Indonesia Census Operations

Introduction

Population census has a strategic value because its result will be useful for development planning in many aspects. In the statistical operation itself, population census data will be a reference for future censuses as well as other surveys. In Indonesia where population registration has not performed as expected, the population census remains the main source of population statistics for development planning, formulation of policies/ programs, and monitoring of progress in achieving development goals.

Since its declaration of independence, Indonesia has carried out four population censuses, i.e., in 1961, 1971, 1980, and 1990. The fifth population census will be carried out in the year 2000. One major decision to be made related to the 2000 Population Census is to publish census data down to the smallest administrative area, i.e., village. In 1990 Population Census, detailed information other than sex could not be published down to the smallest administrative area, because detailed information is collected only in sample. In the census, sample size is only about 5 % of population; hence, it is too small to allow data dissemination less than district (kabupaten/ kotamadya) level. Therefore, to meet many requests for small area statistics, the Central Bureau of Statistics of the Republic of Indonesia (BPS) is planning to collect detailed information of every individual in the 2000 Population Census. Having collected detailed information of every individual in Indonesia, BPS will be able to publish population statistics for smaller areas such as kabupaten (district), kecamatan (sub district), and even village.

The consequence of the decision to collect detailed information of every individual is in the volume to be managed. The volume of questionnaires and numerical data involved in the 2000 Population Census will be enormous. In a country that has more than 200 million population, the census will deal with complex data collection, processing, and dissemination. Therefore, for better results of the 2000 round of population census, UNFPA suggests that national statistical offices develop a realistic

data processing plan in light of experience in previous censuses/ surveys by utilizing the advances in emerging technology information (UNFPA, 1998).

It is a BPS's commitment to always exploit new information technology in achieving its mission. That is why, since 1960s BPS has utilized computers in processing its censuses and surveys data. Moreover, since 1980s BPS has distributed processing capabilities to its provincial branch offices. As a result, distribution of processing can reduce processing cost enormously, because it minimizes the cost of dispatching questionnaires. In addition, the processing time could be reduced, so data can be disseminated in a relatively shorter period.

In recent years, BPS is utilizing Internet facilities, mainly for data dissemination work. However, many Internet users say that many other applications developed under Internet have shown a number of positive results, and it likely more in the future. That is why BPS intend to also utilize Internet in the coming the 2000 Population Census, so this census will also be benefited from this technology.

Realizing the complexity and in anticipating the difficulty faced in carrying out the 2000 Population Census, BPS, under the cooperation with JICA, is currently carrying out a Mini-project called "Looking for appropriate technology for the 2000 Population Census in Indonesia". This project is aimed at searching and studying new appropriate technology in areas of cartographic mapping, data capture, data dissemination and analysis so that the best results could be obtained.

Discussion in this paper will be divided into three areas of census operations, i.e., data collection, processing and dissemination. However, before discussing these three areas, I will describe how we introduce information technology culture in BPS. The success of a new technology is not only to acquire the technology per se, but a plan should oversee whether the technology has been internationally proved in area of statistics and can answer the information requirements. In order to succeed, we should also consider accumulated experiences, available staffs, available infrastructure, staff

skills, and cost aspects (Dekker, 1996). In addition, before closing the discussion, I will also discuss how information technology is able to improve the management of census operation.

Information Technology Culture in BPS

It is understood that the exploitation of a new information technology costs a lot of money. Many institutions have failed to exploit such a technology because human resource in the organization can not support the exploitation. Therefore, BPS concludes that without full participation of the staff, the exploitation of new technology such as information technology will be a failure. In order to achieve the participation of the staff, a new culture that makes them work comfortable by the technology called information technology (IT) culture was introduced in the organization. Since IT is developing very rapidly, the utilization of IT should not be driven solely by technological factor, but human factors should also be taken into account. In fact, human resource is the key component.

The culture of the information technology in BPS has started since early 1960s, where a computer system installed. However, since computer mainframe at that time was installed in an exclusive unit, the technology was only known among a small number of staffs called data processing center people. Mass introduction of the technology was started in 1980s where several PCs were installed in all the province regional offices, as well as in the central office.

As the price of PCs is coming down, BPS has been able to install more computers in the regional offices, both province as well as kabupaten (district)/ kotamadya (municipality) offices. This facility allows BPS to improve its outputs on one hand, and has changed the culture of work on the other hand. By installing computers in every organization unit, BPS has changed several methods of work that is more information technology oriented. In the central office, every unit manager has a computer and he or

she is expected to utilize the facility for supporting his or her jobs. Electronic mails, announcements and other communications are communicated through Internet.

To meet the IT culture requirement, BPS conducts regular training through a permanent in-house training facility. This facility consists of a training center equipped with comprehensive computer training system. In addition to a regular training, BPS performs monthly IT seminars and publishes IT bulletins.

Data Collection and Capture

As the collection stage is the point where data are captured or created, the collection stage should be done accurately and timely. Because, if there is no data produced in this stage, then the following stages, i.e. data processing and data dissemination will have no meaning. The common method in data collection is keying data from questionnaires filled by census enumerators. For years, this keying work was done in the central office, because computer facility was available only in the central office. Since 1980s, population census data has been captured in regional branch offices, as they have been equipped with a number of personal computers. Regional provincial offices entered the census data from questionnaires into computer media, and sent the data to the central office in the form of computer media, instead of in the form of documents.

One major decision related to data capture that is being considered by BPS is to utilize optical mark reader (OMR) and optical character reader/ recognition (OCR) technology. This technology is considered since the system could meet the requirement for collecting data for every individual in the 2000 Population Census. Actually, BPS has already started to utilize OMR facility in 1971 Population Census and 1973 Agriculture Census. But the system was abandoned for several reasons, such as: the high cost of OMR machine, the difficult handling of documents in remote areas, the high cost of a good quality paper in the country, the lack of high precision printing facility in the

country, and the emergence of electronic data entry station (Suharto, 1996).

Fortunately, the progress in the information technology such as artificial intelligence has allowed the current OCR/ OMR beats the drawbacks of the old technology. Therefore, inline with the hope of releasing the result of the 2000 Population Census in a relatively shorter period, BPS is planning to utilize the OCR/ OMR system in the census. As the technology has improved, it is expected that the OCR/ OMR system could be successfully utilized.

To avoid failure of using OCR/ OMR in the 2000 Population Census, an OMR/ OCR study is being carried out. This is a part of mini project carried out by BPS with the cooperation with JICA. By conducting a thorough study of this system, it is expected that the system will be chosen very carefully. Since to run a system, many aspects should be considered. For examples are, among other things, scanner capabilities, OCR/ OMR software, and questionnaire design. A pilot test of the OCR/ OMR questionnaire has just been carried out in the month of July 1998 shows that some features of the questionnaire should be adjusted in order to have better performance.

It is planned that the OCR/ OMR capabilities will be distributed to the regional offices, province or district offices. By having OCR/ OMR capabilities, it is expected that regional offices will carry out data capture work for the 2000 Population Census. As a result, we will not need to dispatch documents to the central processing. Therefore, for the 2000 Population Census, we can eliminate or reduce cost of dispatching more than 200 million documents and cost of storage as well.

Instead of dispatching documents, as a result of capturing data in the regional offices, the regional offices need to dispatch digital data to the central office. Luckily, BPS has experienced with Internet for more than three years. As the technology of Internet has already spread over the country, not only in Jakarta, the Internet services become available in almost all-major cities. The good thing about Internet is that the user pays local telephone connection charges, even though he or she tries to connect to any

Internet server in the world. In the case of fax or telephone, connection fee should be paid according to the distance, whether local, long distance or international calls.

In maximally grasping this benefit, BPS requires all regional branch offices to be connected to the Internet. As a result, currently all provincial offices and a number of district/ municipality offices out of 310 offices have connected to Internet through dialed-up connection. Therefore, as already done with current surveys, the 2000 Population Census data captured in the regional offices will be sent to the central office, and vice versa, by paying local call cost only.

Having all its regional branch offices connected to Internet; BPS expects that the 2000 Population Census will be better than previous censuses, because anything related to the census can be communicated easily with the regional offices. When any problem happened that needs direction from the central office, the use of Internet as a medium of discussion becomes very worthwhile. Moreover, in the last part of this paper I will discuss how the management of population census operation could be done better by using Internet facility.

As the technology of Internet is progressing rapidly, BPS expects that in the data collection area Internet be used not only for delivering information but also for data collection. This means that we can collect survey or census data through Internet. This method is inline with the way people are now behaving, i.e., Internet oriented. For them it is very comfortable to fill in survey or census questionnaire data through Internet, because they can quickly fill in the form and then send the information through Internet without using pen or meeting interviewer. It is expected that in the future censuses/ surveys, BPS will utilize this facility more extensively.

To make use of Internet in collecting data, currently BPS is developing an Internet collection system using electronic Forms on the World Wide Web for statistical activities synopsis. As mentioned in the new Statistics Law No 16,1997, it is stipulated that any institution carrying out a statistical survey should report its synopsis of the survey to

BPS. In the synopsis, the survey conductor should report several informations such as survey title, objectives, period, sample size, method adopted, etc. The blank forms containing the elements and other related information will be placed in BPS Homepage in the Internet to be filled by survey conductors. The availability of this system will ease them in reporting the synopsis, and at the same time, BPS is able to collect synopsis data directly without utilizing any enumerators.

Cartographic Mapping and GIS

As done in the previous censuses, the enumerators of the 2000 Population Census will deal with Enumeration Areas (EAs). An EA is a well-defined area where about 300 households units are located. However, due to high mobility of the population, changing administrative areas and other factors, most of the EAs has changed significantly from initially defined in 1990. For that reasons it is planned that before collecting individual information in the 2000 Population Census, BPS will update all maps of current EAs by redesigning an entirely new system in order to obtain better results. Therefore, EA maps with visible boundaries would be drawn in such a way that they should be still useful for field operations in the coming 5-10 years of household surveys. An ideal number of housing or building units in an EA should be determined such that it would provide a good basis for assigning a relatively similar workload for field workers in urban and rural areas (BPS, 1998).

In addition to EA maps, the field surveyor is expected to also draw village maps where he works. However, it should be annotated that the drawing of the maps in the field is done manually to produce sketch maps. Ideally, the drawing of the maps should use global positioning system (GPS) or other technology, so that the coordinates of the object in the maps are correct. To attain this objective effectively, it is necessary for BPS to equip the 2000 Population Census with a more appropriate mapping technology than that currently used. As the part of the mini project, the study of the mapping is included.

Actually, as the result of the 1990 Population Census, BPS has collected sketch maps for all villages. It should be declared that aside from their limitations, the village maps are valuable assets not only for BPS but also for other government institutions since no other institution has such an extensive maps, not even mappings and survey agency because they have scaled maps only down to the district level. Knowing this assets, BPS has maintained a Geographical Information System (GIS) based on village maps, and higher administrative levels such as Kecamatan (sub district), Kabupaten and Kotamadya. It is noted the difficulties in the process of merging several villages maps into one Kecamatan map. The difficulties occurred because a village map is not properly drawn according to exact coordinate and measurement or not really a scaled map. To have exact coordinate, then we need to use Global Positioning System (GPS).

Maintaining maps in GIS platform will be very useful in the planning stage of a survey or census and also in the dissemination stage. In the planning stage, the availability of maps can be a big help in the design of samples. In dissemination stage, GIS allows BPS to produce more interesting outputs. In fact, many users prefer geographical and colorful presentation because it is easy to understand. Considering the importance of maps under GIS, BPS has determined to carry out another mapping activity in the 2000 Population Census. In this activity, villages and enumeration area will be drawn by the census staff, and then adopted into GIS maps. When the regional offices were equipped with scanners for data capture as planned, the scanners would be also used to capture maps from hard copy into computer digitized form. By using this step, it is hoped that the digitizing using digitizer should be reduced or eliminated.

Data Processing

The processing of population census data has been changed greatly by the implementation of the information technology. In the old days, when mainframe was used, processing of census data was centralized at the central office. At that time, all documents from the regional office were dispatched to the central office. Consequently, the centralized system required a big space to store the documents, a large number of data

editors, coders, entry operators, and validators. In many censuses, BPS had to recruit temporary operators.

The availability of information technology in the form of PCs has changed the way data processing done, from centralized in the central office to be distributed to the regional offices. It is started in the small scale in 1980 Population Census, and more and more in the following censuses. A system that adopts a policy of decentralized processing always faces difficulties in monitoring the processing and ensuring uniform processing. Therefore, it is a “must” to have adequate human resources and infrastructure. That is why BPS has always make efforts to improve the capability of its human resources in the central as well as in the regional offices in addition to the acquirement of the infrastructure.

The availability of BPSnet and Internet will also influence the processing of the 2000 Population Census. By having networking in the central office, i.e., BPSnet, and in the regional offices, real distributed processing can be done. Through this method, different departments can share the same population database at the same time.

Internet will also give major contribution to the process of the 2000 Population Census. By implementing Internet in the processing stage, there will be a significant reduction in the processing cost as well as in processing time because we can reduce document dispatching (if the processing is centralized) or diskette dispatching (if the processing is done in the regional offices).

In case a program gets into some troubles, for example, a bug in the program, or because some functions were not performed well, the staff in the regional office could deliver his questions directly to the program developer in the central office, so the program could be updated and added. All this communication is done through Internet. And when the staff responsible for the program in the central office had done the requirement, he then would be able to send the updated program to regional offices through Internet.

Data Dissemination

In a population census, dissemination stage is very important for users since they can get census results at this stage. That is why BPS determines to exploit information technology maximally so that the user can get the results of the census as soon as possible and in the way the users prefer. Many facilities have been used so that the dissemination is effective, for example: diskette, magnetic tape, and currently Internet. For BPS, data dissemination work is the first area benefited from the utilization of Internet. After the initialization of the connection to Internet in 1995, the first service that BPS provided is maintaining a home page in Internet, the Web site (WWW). The address of this home page is www.bps.go.id. As a statistical office, BPS disseminates both regional and sectoral statistical data including population data in the form of HTML.

Actually, before using Internet, BPS has disseminated data in the form of publication and computer media. But it should be noted that in such services, data seekers have to deliver their requests in written form, by telephone, or in person in BPS. On the other hand, data dissemination using Internet does not require a lot of effort from data seekers, because data is in front of him in the computer. It does not matter where he is, whether abroad, in the office or at home. As long as he can connect to the Internet, he can get the information available in the homepage.

The effective use of the Internet in data dissemination stage can be seen in Table 1. This table shows the number of hits, file accesses, data sent, and number of sites accessed to the BPS' home page. It can be seen from the table that, in one year (July 97 to June 98), there are 1,042,196 hits to the home pages, and 781,457 files accessed. Accesses come from 43,371 sites and they have already transmitted 4,154,419 byte data.

The users accessed the home page is not only from Indonesia but also from almost all countries in the world. And in fact, the biggest number of users is from the United States and is more than the number of users in Indonesia. The ten biggest numbers of users are from the United States, Japan, Australia, Canada, Netherlands, United

Kingdom, Germany, France, Thailand and Malaysia.

Table 1. Number of hits, files accessed, Kbytes sent, and Sites accessed to BPS' Home Page by Month in July 1997 to June 1998

Month	Hits	Files	Sites	K bytes sent
Jul 97	155091	89121	2742	440108
Aug 97	108945	62261	1958	305396
Sep 97	160378	105283	3249	539291
Oct 97	103933	60044	1761	304762
Nov 97	75003	52621	2950	272217
Dec 97	58274	46709	2960	237827
Jan 98	58519	57791	3758	300185
Feb 98	52362	50343	4139	264009
March 98	37981	35940	4521	206879
April 98	72684	69837	5114	372962
May 98	81921	78381	5676	472633
June 98	77105	73126	4543	438150
TOTAL	1042196	781457	43371	4154419
Average	86849.67	65121.42	3614.25	350476

Source: http-analyze 1.9 Report

The homepage of BPS is initially defined and prepared by the headquarter itself. However, as the regional offices become more familiar with the Internet facilities, they are also invited to prepare the Internet by themselves. In this case, they can prepare pages specialty with the region. For that reason, each regional office is now asked to maintain their homepages.

One facility that is included in the home page system is the capability to

communicate to BPS webmaster and BPS' Data Dissemination Division. This facility allows outside users to deliver their comments, questions, and also data requests through Internet. And in fact, many data seekers have utilized this facility. An average of two requests received by the Webmaster and the Data Dissemination Division every day.

One other method that is being considered by BPS in disseminating population census data is using CD-ROM. This method is more acceptable for the users since CD-ROM driver become a common part of a computer system, and is secured. For that reason several CD-ROM writers have been acquired and software accompanied the data is being developed. It is planned that the 2000 Population Census could be disseminated in any way a user wants.

Management of the 2000 Population Census

Quality check should be an integral part of the census data collection and processing to ensure that census data will be at an acceptable level of quality. That is why from the beginning BPS has determined to adopt the information technology (IT) in managing the 2000 Population Census operations. By having used the IT, top managers can monitor the progress of the activities. Where decentralized operation will be used in the 2000 Population Census, monitoring and uniform processing becomes a challenging task for BPS. The decision on the number of processing centers should be made not solely on the basis of hardware availability alone, but also other aspects such as human resources as well as the monitoring methods.

Actually, the implementation of the IT in the management of survey activities has been used by BPS since 1986. The system called KABSIS stands for Kabupaten Statistical Information System has been used in monitoring the state of the enumeration for three surveys, i.e., the Annual survey on medium and large scale industry, the Quarterly survey on medium and large scale industry, and the Wage survey. By using KABSIS, the head of the Kabupaten/ Kotamadya Regional Office can know how far the collection of data has been taken place and the promptness of the collection. This will be

an early warning system for him to know which establishments are difficult to enumerate and the performance of the enumerators. Thus, he can take the necessary steps to make sure that all the establishments in his working area submit their data to BPS. (BPS, 1986).

Therefore, BPS is contemplating the monitoring process of the entire the 2000 Population Census operation in the central as well as in the regional offices by utilizing information technology, not just using conventional methods such as logbooks and planning boards. A census requires careful and detailed planning management, because the census may be unique in the sense that no activity of similar size and complexity has been conducted regularly since the census taking is only once in ten years. A census has to observe strict deadlines so a postponement could be avoided. Therefore, some software especially for census management such as the Census Tracking System, a special facility of IMPS, is also considered to be utilized.

BPS is also planning to exploit new IT in managing the 2000 Population Census. The use of Internet in carrying out the 2000 Population Census should be explored carefully in advance. Management is one area that can be benefited from the use of Internet. As we know, a large-scale undertaking like Population Census would consist of a number of activities related closely to each other and the voluminous work of the census would require a large number of staff and part time enumerators. So it is very important for the project manager as well as the high-ranking officials of BPS to know the state of the development of the project at any point of time. And with the support of the Internet, BPS is able to manage the progress as well as the obstacles to be solved in the field.

For that reason, there will be a system that allows every regional office to enter the field work status into the system, and send the data to the central office through Internet. The system then process the data automatically and then provide summary status of the Population Census undertaking at a particular day. Looking at this figure, the top manager will immediately aware the status of the work in the field, and observe

the probable deviation compared to the designed and scheduled activities.

As BPS has planned to capture data using OCR/ OMR machine and data cleaning in the regional offices, it is expected that the cleaned data could be sent to the central office through the Internet to reduce the time lag in the data processing.

However, the use of Internet is needed not only for data communication but also for consultation. In the stage of coding and data cleaning, there will be a lot of questions and answers from and to the regional offices. The central office and the regional offices perform these activities through the Internet.

Concluding Remarks

For BPS, the exploitation of the information technology in conducting census operation is not new. In fact, since 1960s until now, the availability of the information technology facilities in BPS has been mainly to support censuses. The long experience in the information technology area has changed BPS' practices into IT oriented, and BPS always tries to grasp the best of the technology in order to improve its census or surveys results. Since IT is developing very rapidly, the design of a data processing system should not be driven solely by technological factors. Other factors, such as human and financial resources, and training needs, should be taken into account. For BPS, human resource is the key component to the success of information technology application, so that the development of human resource in IT is done continuously.

It can be seen that the way of data processing work has been changed in terms of data collection, data processing and data dissemination. Inline with the intention to collect detailed information from every individual in Indonesia, BPS plans to utilize OCR/ OMR systems in the 2000 Population Census. Therefore, a through study is being carried out so that the system will be really successful. Also in the collection stage, Enumeration Area (EA) and village maps will always be a basic tool for carrying a

population census as well as other censuses/ surveys. Therefore, to preserve the maps and to make it easily accessible, BPS employs geographical information systems (GIS) and CD-ROM technologies. In turns, the availability of the maps in GIS will contribute not only in the collection stage, but also in the dissemination stage where many geographical statistical presentations can be produced.

One major advantage of using information technology is a reduction of processing time. In the old days, processing time of population census could be more than two years, and BPS is optimistic that the time can be reduced to a relatively shorter period based on the availability of technologies such as OCR/ OMR system, networking and Internet.

An Internet facility has also one of the major factors in performing effective data dissemination. In fact, for years BPS has utilized this facility, and records show that many users welcome this facilities. Therefore, Internet will also be one of the major approaches in disseminating the 2000 Population Census results as well as in computer media such as CD-ROMs, diskettes, magnetic tape, etc.

Since the operations of population census deals with many enumerators, supervisors, data entry operators, documents and financial resources, the management of the census should be carefully carried out with a little more sophistication. Understanding this complexity, BPS is contemplating to exploit information technology in supporting management. The support will also use Internet, which allows delivering regional reports to the central office regularly. And the top manager could monitor the whole state of the operation every time as necessary.

ooOoo

Reference

- BPS, Pedoman Pelaksanaan KABSIS: Sistem Pemantauan Pemasukan Dokumen di Tingkat Kabupaten/ Kotamadya, Badan Pusat Statistik, Jakarta, 1996
- BPS, Sensus Penduduk 2000: Pedoman Kantor Statistik Propinsi dan Kabupaten/ Kotamadya, Badan Pusat Statistik, Jakarta, 1998
- Dekker, Arij, Innovative Computer Methods for Demographic Enquiries and the Dissemination of Population Information, Proceedings of the Expert Group Meeting in Innovations Techniques for Population Census and Large-Scale Demographic Surveys, Nidi/ UNFPA, 1996
- Suharto, Sam, Innovative Techniques in the 1990 Round of Population and Housing Censuses, Possibilities for the Future, in Technical Notes on Population Statistics and Information Systems, issued by Statistical Division of the Department for Economic and Social Information and Policy Analysis (DESIPA), June 1996
- UNFPA/CST, Census and Survey Data Processing in East and South-East Asia in the 1990s: Issues and Lessons Learned, Occasional Paper Series No. 6, UNFPA/CST, Bangkok, 1998