

United Nations Statistical Institute for Asia and the Pacific

Sixth Management Seminar
for the Heads of National Statistical Offices in Asia and the Pacific
28 - 30 May 2007, Hong Kong, China

**Managing Statistical Development and Information Technology for
National Statistical Offices**

Country Paper: Macao, China

Meng Chao IEONG
Acting Deputy Director
Statistics and Census Service
Macao, China

1. Background of ICT Development

The Statistics and Census Service (DSEC) is the primary government agency to produce official statistics in Macao. Back to the late 70s, DSEC was one of the first government agencies to use mainframe computer (NCR and IBM) to run the business. Though the processing power and the storage capacity of the “mainframes” was very little compare to the systems today, the mainframe computers made a good contribution in shorten the delivery time of survey results at that time. The Population Census 1980 was the first population census done by computers in Macao history.

In the early 80s, DSEC started using PCs in statistics compilation. The software packages used were mainly SPSS, SAS and some tailor-made application written in Excel, Basic and COBOL.

The rapid development of PCs sooner became the major computing platform in DSEC. Local Area Network (LAN) was established in early 90s, and Client/Server computing was later employed. In 2000, in face of the Y2K problems, DSEC took the opportunity to “downsize” the computing systems, and the IBM host computer was retired since then.

2. The ICT Development Team

An experienced in-house development team is responsible for all the ICT projects in DSEC. The team is composited with 14 programmers and 4 supporting staff. This team size is very small in consideration of large numbers of surveys projects. The team can be further divided into small groups:

General Survey Projects Group. This group is responsible for the development of application for the general survey projects. Group members are mainly experienced programmers in statistics computation. Most of them grew in the early computing age, they are familiar in procedural programming languages and statistics software packages such as SPSS and SAS. This group can be seen as the foundation of the ICT team which ensures the daily operation of the survey projects.

New Application Projects Group. This group is responsible for the development of the new ICT applications, such as CATI, OCR and PDA applications. Group members are usually young professionals which has fresh ideas and advanced I.T. skills. The output of this group will be introduced to the rest of the team thus surveys projects might make use of these new applications.

The Internet Group. This group is responsible for the development of the Internet related applications. As the penetration of the Internet in Macao is growing fast, there are huge opportunities of using the Internet in statistics compilation. Online data submission, online statistics databases queries, web reporting are some areas that the group will works on.

Users Support Group. This group is responsible for the daily supports for the end users. Their jobs include hardware installation, software installation, security settings, virus prevention etc.

3. Project Priorities and ICT Budgets

Efficiency gain is always the first consideration in project priorities and ICT budgets. Take 2001 population census as an example, the census period was fixed in 13 days, the computer systems must support the operation with highest efficiency possible, so a highly integrated and sophisticated computer system was proposed and built (refer to Point 5 for details of this project.).

In Macao, all government budgets are prepared and go through the approval process in the last quarter of the preceding year. The ICT budget is approved in a total amount. Once it is approved, the agency can decide the uses of the ICT budget among different projects.

4. ICT Training and Government Support

In Macao there is a government agency responsible for all the training activities, including ICT trainings. The ICT trainings are usually general purpose, such as Word, Excel, Flash etc. Some programming courses have also been organized by the agency, such as Java and .Net. However, these training courses are usually not technologically deep enough for our I.T. staff to advance their specific I.T. skill. These create a need to conduct in-house ICT trainings in DSEC.

The in-house ICT trainings are usually more specific to one problem domain, such as a specific technology or programming language. The trainer could be our own staff which had studied the subject matter for certain time and had accumulated reasonable working experience on it. If internal trainer is not available, the training course might be consigned to an external vendor or consultant.

As ICT is changing so fast, it is not quite possible to have all I.T. staff trained in the same level. We encourage our staff to invest their time in the specific subject which he/she is interested in, and try to arrange appropriate jobs for them to apply what they learnt. This approach works well in DSEC as individuals are motivated to learn their favour skill, and their knowledge can contribute to the works, results in a win-win situation.

Worth mentioned that individual knowledge or experiences are encouraged to be shared with other colleagues. This not only help sharing knowledge but also building team spirit. We usually conduct knowledge exchange sessions on Friday afternoon. The week-end mood makes the atmospheres more relax. If possible, drinks and cakes could also be served. It helps.

5. Recent Completed Project: **By-Census 2006**

The computer systems built for the By-Census 2006 was a showcase of all the existing ICT applications employed by DSEC. The system is designed in a two-tiers architecture: the head office system (Backend system) is responsible for central operation such as OCR, CATI and progress monitoring. The branch offices system (Frontend system) is responsible for the field works management. (See Appendix for the details of the network architecture.)

The following are the features of the By-Census system:

- a. **Broadband Internet Connection:** In order to provide a highly efficient and smooth operation between backend and frontend systems, broadband Internet were employed to form the Virtual Private Network (VPNs) for this operation. All data and images of questionnaire were sent through the VPNs.
- b. **Onsite Questionnaire Image Capturing:** In the past, errors were usually found in the questionnaires - after the Census period. It was a time consuming process to clean these questionnaires errors because households might needed to be re-visited (or contacted by telephone) in order to find the correct answers. The later the errors were found, the more difficult to correct the errors.

To tackle this problem, the Macao By-Census 2006 used an ambitious approach to verify the questionnaires once it was returned by the enumerators. The process includes the following steps:

Step-1: Questionnaires were scanned in the branch offices;

Step-2: Scanned images were sent to the head office through VPNs;

Step-3: OCR processing (including manual verification) was performed in backend system installed in the head office;

Step-4: Validation processing was performed immediately in the backend system to find out the errors;

Step-5: The results of the validation process were sent back to the branch offices;

Step-6: If the validation result shows errors in the questionnaire, the enumerator will be informed and responsible to clean the errors, and re-submit the questionnaire (back to Step-1);

Step-7: If the questionnaire is clean (no errors found), it would be filed and sent to the head office at the end of the Census operation.

The above processing cycle were scheduled to be completed on a daily bases, so the enumerators could see the errors (if any) the next day he/she submitted the questionnaires, and immediately correct the errors by re-contacting the households.

- c. **Decentralized Scanning and Centralized OCR:** questionnaires were scanned in census branch offices and OCR processing (including manual verification) were performed in central office, this approach eliminated the need to deliver the paper questionnaires between offices, and the resources (people and computers) are fully

utilized. Furthermore, coding for open questions (like occupations and positions) was centralized in head office. This not only maximized resource utilization, but also helped quality control.

Cons: re-submitted questionnaires spent redundant resources in OCR, coding, and validation processes.

- d. **Integrated CATI Support:** CATI system serves three functions in the By-Census 2006 operation: (1) though it was not recommended, households could answer the questionnaire through CATI; (2) household could make appointment of the home visits through the CATI, and (3) general inquiries.

It was welcome by the households to make appointment with our enumerator for the visits. This not only help the household to have the enumerator visited on their preferred time slots, but also help the enumerators finished their job faster.

- e. **Integrated SMS Support:** Census is a dynamic project, smooth messaging plays an important role in guarantee a successful operation. The By-Census 2006 employed a simple but effective SMS approach to establish the communication between the offices and the field enumerators. The SMS were generated by the computer system to remind the enumerator of specific events, such as an urgent appointment requested by household.
- f. **Total Progress Management:** Visit results were inputted in branch offices everyday during census period, managers in both branch offices and head office had total monitoring on the teams' progresses. Thi helped managers to make in time decisions such as re-allocation of resources in different branches or different operations.

The By-Census 2006 demonstrated a good use of ICT in census project. The combined use of different technologies made the data collection more efficient, improved overall data quality, and resulted in a faster release of census results.

6. Recent Developing Project: **Time Series Database**

In face of huge volume of statistic data ever published in Macao, we are building a centralized statistical database to store and disseminate the data to the general public. The idea is to have a central repository to store whatever statistical indicator that we ever compiled, and allow the users to query this database through Internet. In this sense, the users might perform the analysis they want.

The following table presents a simplified structure of this Time Series Database:

Indicator Tree	1970	1971	1972	2005	2006	2007
Demographic								

Population								
Male								
Female								
Birth								
Death								
Employment								
Employment Rate								
Unemployment Rate								
...								
...								

The Indicator Tree can be updated and expanded as required, and the periods can also be added.

In the primary plan, this database will provide the following functions:

- a. Web reporting service: user can build their own statistics report in the way they want through the Web browser. He/she can select the indicators, compare their values and generate their reports. The reports can be exported to Excel or PDF formats.
- b. Mobile Web Pages: a simple and efficient web page will be built to serve the mobile Internet user, such as PDAs and smart phones users.

7. Conclusion

DSEC deeply believes that ICT plays a vital role in statistics compilation and dissemination, and believes that it is one of the key elements to assure quality statistics data. The following are some areas of ICT application that DSEC is working on:

Dynamic Data Dissemination on Web: With the sophistication of Web technology, much richer contents can be presented on Web. Dynamic contents such as GIS or Web reporting can be easily implemented on Web.

Mobile Computing: The development of wireless technology and the smart handheld devices open new possibilities in statistics compilation and data dissemination. Data collection by PDAs, special homepages designed for the small screen of handheld devices, enumerators field works with GPS assists, are some areas we are working on.

Security Protection: The advancement of ICT technology, especially the Internet technology, also creates a challenge for the NSOs to protect the data. Prevention from the attacks of virus, spyware, and other malicious softwares are the top mission for the I.T. team of DSEC.

Appendix: Network Architecture Diagram of By-census 2006 of Macao

