

IV. Online Data Collection for Statistical Surveys in the Government of Japan

Bunzo HIRAI

Statistical Standards Department

Statistics Bureau

Ministry of Public Management, Home Affairs,

Posts and Telecommunications

JAPAN

1. Government Policy for Online Data Collection

(1) The New Strategies for Government Statistical Services for the Coming Decade

The New Strategies for Government Statistical Services for the Coming Decade, a comprehensive policy vision on official statistics in the Government of Japan submitted by the Statistical Council in March 1995¹, first proposed the introduction of online data collection for statistical surveys as follows.

Chapter 5. Efficient Implementation and Quality Improvement of Statistical Surveys

1. Statistical Surveys in the Information Age

1) Capturing Statistical Data through Magnetic Media or/and Telecommunication lines

The development of information and communication technologies makes it possible for information and communication equipment such as computer hardware to be disseminated, for telecommunications networks to expand, for personal computer networks to develop and for telecommunication and software business to prosper. Under these circumstances, enterprises and establishments are equipping themselves with electronic instruments. These enterprises and establishments are not only reducing the cost of business administration and the time used in data exchange by computer but also starting to connect their computers to business partners' to order goods or receive orders (electronic data interchange), hence expanding computer applications.

As most of enterprises and establishments are equipped with computers, magnetic media or telecommunication lines need to be used in statistical data collection so as to carry out statistical surveys efficiently and lower the response burden.

a. Capturing Data through Magnetic Media and/or Telecommunication Lines

¹ The Statistical Council is an advisory body of the Statistical Standards Department, Statistical Bureau, Ministry of Public Management, Home Affairs, Posts and Telecommunications (MPHPT) (Management and Coordination Agency up to January 2001), composed of statistical experts.

In the case of monthly surveys which collect data on the same items frequently, if respondents handle their transaction data by computer, once software is developed to translate data already recorded on respondents into the data requested by a survey, data for that survey will be processed almost automatically using the respondents' original data. As a result, the response burden will decrease dramatically. Enterprises and establishments are therefore demanding that government statistical offices exploit magnetic media such as floppy disks or telecommunication lines such as computer networks in collecting statistical data.

In addition, data collection by magnetic media and/or telecommunication lines improves the efficiency of field operations by reducing the cost of checking the contents of questionnaires and decreasing the number of enumerator surveys. Furthermore, such methods of data collection make the arrangement of questionnaires more efficient and reduce the volume of data input, so that the period necessary for data processing is shortened and timelier dissemination of data becomes possible.

As can be drawn from the above, we expect to see benefits in all statistical survey processes, that is, from the survey design stage to the data dissemination stage, through the introduction of data collection via magnetic media and/or telecommunication lines.

Therefore, data for surveys on enterprises and/or establishments should be collected using telecommunication lines such as computer networks and/or magnetic media such as floppy disks, starting with surveys conducted frequently (monthly or quarterly).

On the other hand, as dissemination levels of information and communication equipment in households are relatively low compared with those of enterprises and establishments, at present it would seem difficult to collect data from households or individuals by magnetic media and/or telecommunication lines. Rather, it is desirable to continue to investigate the possibilities of data collection by magnetic media and/or telecommunication lines from households or individuals, keeping an eye on changes in the dissemination levels of information and communication equipment.

As we can expect the appearance of cheaper hardware of higher capacity and communication lines of larger volume with cheaper costs, we have to examine the use of information and communication technologies questionnaire distribution part of the survey process.

b. Points to be Kept in Mind in Collecting Statistical Data Using Magnetic Media and/or Telecommunication Lines

When the government collects statistical data using telecommunication lines, the unfortunate possibilities arise of data being sent to the wrong number and the disappearance or leakage of data by illegal access for the purpose of data theft. To avoid such possibilities, it is recommended that such measures as setting ID codes and/or passwords, the strict management of these and the division of computer systems into data transmission systems and internal processing system should be taken whenever possible.

Furthermore, because information and communication equipment, data formats and communication protocols differ between respondents, it is desirable to study setting up magnetic media standards and communication protocols to be used in statistical surveys, considering the trend toward standardization of hardware and software and increasing utilization of Electronic Data Interchange.

(2) The Promotion of Statistical Surveys via Electronic Measures

In March 1999, directors of statistical divisions of ministries and agencies reached an agreement named the Promotion of Statistical Surveys via Electronic Measures.

The agreement stipulates that the purposes of utilizing electronic measures is i) reducing the response burden by streamlining respondents' time and labor to entry data, ii) efficient implementation of statistical surveys by rationalizing data check at survey-conducting entities such as local governments, and iii) prompt publication of survey results by streamlining data entry by the offices which conduct surveys.

The coverage of statistical surveys in this agreement is for the time being those which collect the same items so frequently as monthly from establishments and/or enterprises because the effect of the measures seems to be realized relatively early in such surveys on account of the diffusion levels of information and communication equipments.

The method to be used is basically online one such as Internet, paying attention to smooth collection of statistical data and data security.

In order to reduce the response burden, the offices which conduct online surveys should make efforts to simplify and streamline data entry workload of respondents when designing and developing surveys. In addition, reporters should be allowed to choose alternative methods to report such as reporting with conventional questionnaires.

2. Examples of Online Data Collection in the Government of Japan

(1) The New-Generation Statistical System of the Ministry of Economy, Trade and Industry (METI)

The Ministry of Economy, Trade and Industry developed the New-Generation Statistical System, a system to process the reports of statistical surveys online utilizing Internet, and started the operation of the System in January 2000.

The System aims not only to reduce the response burden using online reporting via Internet but also to improve services (data dissemination etc.), to secure promptness and to reduce the costs by processing data electronically.

The coverage of surveys are current surveys (monthly surveys) such as Current Production

Statistics Survey of METI, Current Survey of Non-Ferrous Metal Supply and Demand, Current Survey of Petroleum Products Supply and Demand, Current Survey of Commerce, Statistical Survey of Selected Services Industries and Survey of Energy Consumption in Commerce and Manufacturing.

The System also allows the Ministry to produce electronically some indices processed from the above-mentioned statistical survey results such as Indices of Mining and Industrial Production, Shipments, Inventories and Inventory Rates, Indices of Tertiary Industrial Activity and Indices of Commercial Sales Value by Type of Business.

The System adopted a security system in which survey respondents are requested to prepare their system that supports the SSL (Secure Socket Layer) protocol and to register electronic certification, and thus information is encrypted and transmitted based on SSL prepared by METI. Also, in respect of the procedures for electronic certification to utilize the SSL protocol, an institution for examinations and registrations, and that for issuing and managing certificates are separated. As for functions relating to examinations and registrations, METI examines survey respondents, and functions relating to the issuance and management of certificates are consigned to and implemented by private enterprises.

(2) The Online Application System of the Ministry of Land, Infrastructure and Transport (MLIT)

The Online Application System of MLIT focuses mainly on application and reporting for administrative action such as permission and licensing, but also covers statistical reporting. The system started its operation in FY 2001

Since the System covers both administrative procedures and statistical surveys, the software to be used by survey respondents are common to that to be used by applicants for administrative action. In order to respond the standardization trend of electronic data interchange (EDI) in business, the System adapted XML format. For the security on the Internet, the certification technology of the public key infrastructure (PKI) is used in the System.

At present the surveys covered by the System is limited only to those for transport statistics such as Survey on Air Transport (monthly), Survey on Railway Transport (monthly) and Survey of the Trend of Equipment Investment By Transport Enterprises (annual), all of which are approved statistical surveys based on the Statistical Reports Coordination Law. MLIT now runs trial to implement online report collection for surveys for statistics designated based on the Statistics Law such as Survey on Current Rolling Stock Production and Coastal Vessel Transport Survey (both monthly).

(3) Online Surveys by the Ministry of Health, Labor and Welfare (MHLW)

MHLW started online data collection for statistical surveys via Internet in FY 2001. Currently reports for Monthly Labor Survey (designated statistical survey, monthly) and Survey on Labor Economy Trends (approved statistical survey, quarterly) are collected online. Both of them are establishment surveys.

3. Conclusion

The diffusion of high-performance PCs and Internet in the last decades has enabled us to develop online data collection system rather easily and inexpensively. The broadband technology enhanced the speed of data transfer. The development of SSL and PKI has ensured secure environment for such collection.

So far the introduction of online data collection is almost limited to frequently reported enterprise and establishment surveys because information and communication equipments are more popular in enterprises and establishments and the effect of reduction of reporting burdens is larger in frequent surveys with the same survey items, which are usually found in enterprise and establishment surveys.

The future issue is the introduction of online data collection to surveys on the households and individuals. This would depend on the diffusion levels of information and communications equipments and broadband connections in the households.