New Generation Statistics System: METI's Approach to Internet Statistical Data Collection and Analysis

Kazuyuki MOTOHASHI[1]

Ministry of Economy, Trade and Industry, Japan

Abstract

This paper presents features and effects of the New Generation Statistics System at METI, an innovative statistical system which can handle various statistical activities, such as online data collection, compilation of secondary processed statistics based on such data and provision of feedback of data to respondents, by means of Internet technology. This paper also provides some background information behind this system, such as GOJ's e-government initiative, and future perspectives of the system for improving the online survey collection rate.

Key words:

Online data collection, security of data transaction, METI's monthly Industrial Statistics

1. Introduction

In January 2000, the Japanese Ministry of International Trade and Industry (now called the Ministry of Economy, Trade and Industry) started to operate the New Generation Statistics System (NGSS), a statistical system whereby a series of statistical operations, such as the collection of survey results, their analysis and the feedback of the data to respondents, is conducted by means of Internet technology. METI began designing the system in 1995 and perfected it after about five years of system development and experimental use. The innovative NGSS is the first large-scale, online statistical system in Japan, and it can be used for all the METI monthly industrial surveys, and for compiling economic indicators such as the Indices of Industrial Production (IIP).

Below we look at some of the background to the development of NGSS.

(1) E-Government Initiative by GOJ

Along with aiming for greater efficiency and improved quality of governmental services and working towards a reduction of the burden on the people of governmental procedures, the Government of Japan plans to conduct all administrative procedures in such a way that

---

[1] Executive Deputy Director, Research and Statistics Department, Ministry of Economy, Trade and Industry (METI), 1-3-1 Kasmigaseki, Chiyoda-ku, Tokyo, Japan, e-mail: motohashi-kazuyuki@meti.go.jp. Views expressed here are author's own and do not represent those of his organization's.

applications can be made online, in order to streamline a fundamental data base, in principle by the year 2003. Among these e-government initiatives, one important field is the online receipt of statistical survey reports. NGSS was developed in advance of this initiative, and is receiving a lot of attention from each statistical section in the government, as the best practice for the online receipt of statistical surveys.

Moreover, the e-government initiative does not entail simply conducting all kinds of current government procedural operations by electronic means, but rather reflects a trend to review business styles so that better-quality administrative services can be provided for the people. NGSS is not only for the purpose of collecting survey results online; its greater significance lies in the fact that it can handle various statistical activities, such as compiling secondary processed statistics by using survey data collected online and providing feedback about such data to respondents, in a uniform and consistent manner.

(2) Improving the statistics and reducing the load on respondents at a time

As economic activities grow more complex, through means such as the development of outsourcing and the integration of product technology, it is increasingly necessary for statistics to be more thorough in order to obtain greater accuracy in measuring economic activities. However, at the same time it becomes increasingly difficult to count on the cooperation of respondents in statistical surveys. In order to cope with such circumstances, it is important to reduce as much as possible the load for each respondent, as well as to make the best possible use of the collected data. NGSS makes it possible for the complied data to be released promptly, through rapid processing, and to be analyzed from many different viewpoints. NGSS has been developed with detailed consideration given to reducing the data entry load as much as possible for each survey body. Also, processing the data that has been provided online as it is ensures the early release of the statistical data.

(3) Advancement of Information Technology

Finally, for the realization of NGSS, using network technology for client server systems, and taking advantage of the diffusion of the Internet, IT plays a major role. While traditional systems mainly conducted batch processing through mainframe architecture, NGSS is designed to meet the needs of end users more flexibly. Moreover, there is a need to set up user-side data infrastructure for electronic data collection systems. In that sense, the recent explosive spread of the Internet is highly significant.

2. Features of NGSS

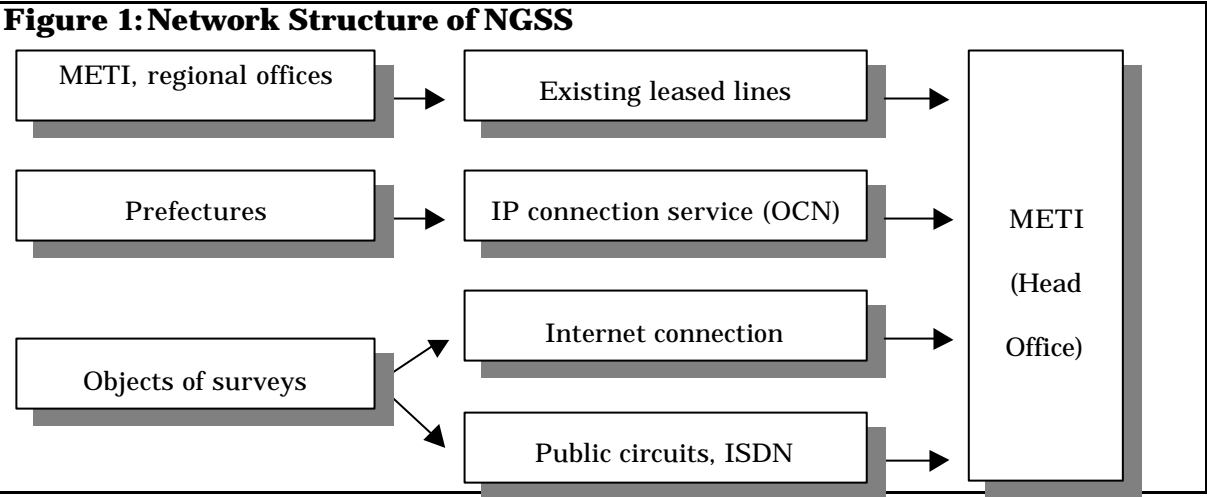*(1) Statistical surveys covered by NGSS*

The statistical surveys conducted by METI can be divided into the large-scale census statistics that are carried out at the most once a year or once every several years, and the monthly or

quarterly current statistical surveys. NGSS targets such current statistical surveys, in other words, current statistics on production from manufacturing industries, current commercial statistics from the wholesale/retailing industry, and special service industry statistics from the service industry, and has made online statistical reporting possible for more than 190 kinds of questionnaire surveys. In addition, secondary statistics such as the Indices of Industrial Production, the Indices of Tertiary Industrial Activities and the Quick Estimate for Commerce Sales, which are prepared on the basis of these statistics, are also calculated using NGSS. Details of statistics processed by NGSS are provided in Appendix 1. Furthermore, tabulated data is passed on to those taking the survey, as well as the prefectures of Japan and METI's regional offices, the supporting organizations of these current statistical surveys, and this circulation of data can also be carried out by NGSS. A diagram of NGSS's flow of transactions is provided in Appendix 2.

(2) Wide-area-network statistical system

The statistical surveys to be covered by NGSS are not conducted by direct interaction between survey respondents and METI, but involve METI's regional branches and Japan's prefectures. Therefore, NGSS is constructed between survey respondents, prefectures, and METI's regional branches and headquarters, and facilitates data collection, examination, correction, feedback, provision, etc., with this nation-wide network infrastructure.

As is shown in Fig.1, this wide-area network consists of existing leased lines between METI's headquarters and regional branches, OCN (IP connection service provided by NTT) between METI and prefectures, and various types of communication systems between survey respondents and METI. As for survey respondents, there are two options for communicating with METI: accessing the password-controlled Web-page through the Internet, or   connecting to the NGSS's access point via dial-up through public telecommunication lines, such as ISDN. It should be noted that telephone and ISP communication fees are paid by METI, because this is an obligatory statistical survey.

Figure 1: Network Structure of NGSS

METI, regional offices	→	Existing leased lines	→	
Prefectures	→	IP connection service (OCN)	→	METI (Head Office)
Objects of surveys	→	Internet connection	→	
	→	Public circuits, ISDN	→	

(3) Up-to-date security system

Due to the confidential nature of individual statistical survey information, maintaining a secure environment in data exchange is critical to operating NGSS. Wrongdoings on the network could include wire-tapping, tampering, impersonation and misrepresentation, ex-post rejections, software virus infections, etc. Leakage and wire-tapping refer to the leakage of internal information through the unauthorized entry of an outsider, or the theft of information by a third party over the Internet, which can be dealt with by using measures including access control based on the establishment of the right of access, and the encryption of data, by which a sender can scramble it. Tampering refers to a change in information by a third party on the Internet. In order to counter this problem, a digital signature, the equivalent of a signature and seal in the case of documents, can be used.

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

In addition, the problem of software viruses must also be dealt with. Especially in the development of the system, a great deal of energy was directed into creating countermeasures to solve this problem. This was due to our conclusion that the securing of the reliability of the new system is an important point in its promotion, as there is the possibility that a software virus could infect not only the Ministry of Economy, Trade and Industry, but respondents as well.

3. Effects of NGSS

(1) Reduction of burdens on survey respondents

NGSS is designed to reduce the statistical burdens on survey respondents. The following are some related features of NGSS.

If a respondent has data in computerized form, she/he does not need to fill in survey sheets, and only has to copy and paste the data into the survey format.

Fixed information, such as the establishment number, etc., can be automatically input and

> ➤ If fixed information, such as establishment code, questionnaire code, etc., are registered in the directory master, survey data are automatically provided in a survey sheet.

Simple checking of inputted data is possible

> ➤ Survey data submitted on an online basis goes through a simplified checking program at the time of storage in the METI computer. Respondents, after transferring the survey data, will be able to check and modify their mistakes in choosing sections, based on the possible error signs coming from the checking programs.

Data submission by a headquarters with several responding establishments

> ➤ Since METI's industrial activity surveys are conducted at the establishment level, a large firm may have several responding establishment. NGSS enables head offices to tabulate survey data for each establishment, and to submit them online at the same time.

(2) Improvement of data feedback services

Taking advantage of its nationwide network infrastructure, NGSS makes it possible to improve the following data feedback services to respondents as well as to regional supporting organizations for statistical surveys.

Data feedback to respondents

> ➤ Respondents can refer to and download, online, their own individual questionnaire data and nationwide tabulated values with respect to survey items reported. Time-series data can go back to1997.

Data feedback to regional supporting organizations..

> ➤ For supporting organizations, i.e., prefectures and METI's regional branches, online operation makes it possible for prefectures to refer to and download nationwide tabulated values and individual questionnaire data, and at the same time, provides a tool for preparing tabulated values in local areas.

Improvement of data dissemination through the Internet

> ➤ It is obvious that published data can be disseminated efficiently through the Internet. Before NGSS was introduced, detailed data dissemination had been done through regional supporting organizations in the off-line mode. Therefore, Internet dissemination contributes to the elimination of regional differences in obtaining published data, etc.

(3) Improvement of promptness

NGSS contributes to the early publication of data due to shortening the process time in the following steps from data collection to data publication. If we can make online submission of almost all survey data, the timing of the data publication can be one week earlier than the

present date.

Data collection
- ➢ Since NGSS enables respondents to fill in survey sheets more easily and quickly, the time spent by respondents can be shortened. In addition, online submission can save the time spent in mailing services.
- ➢ In addition, questionnaire data are processed through a checking program to see if there are simple errors, and examiners at METI conduct a more detailed check. This process is also shortened by the automated data transfer system. In addition, questionnaire data can be examined each time it arrives.

Data processing, editing and publication
- ➢ One can easily do data processing by extracting the necessary data from an Oracle database, and editing them with Windows-based software packages, such as ACCESS and EXCEL.
- ➢ Time spent in index calculation is shortened, by making the transfer of data of primary and secondary statistics systems more efficient.

(4) *Improvement of economic efficiency*

Improve efficiency in data processing works by end-user computing
- ➢ Under the previous system, data processing was conducted in a batch style, after all data were collected. In contrast, NGSS makes it possible for end-user computing by each METI staff member in an interactive style. Survey data can be processed in a decentralized and flexible way.

Reduction of expenses for data input, etc.
- ➢ As data entry for questionnaires will become unnecessary, expenses for data entry staff will be reduced.

4. Operational Status of the System and Future Issues

(1) Improvement of the Rate of Online Submission

As of June 2001, a year and a half since NGSS came into operation in January 2000, the rate of online submissions based on the questionnaire survey was approximately 13%. While it can be said that introduction of the system is steadily making progress, the pace of the rate of increase is slowing down, as is shown in Fig. 2, and it is necessary to further promote dissemination activities in the future. As an activity for dissemination and publicity, along with setting up an NGSS promotional page on the homepage of METI, since 1998 we have been holding information sessions for employers through related business groups, and have prepared and distributed a leaflet promoting NGSS.

(Fig. 2)

To assess users' perceptions of NGSS, we conducted a questionnaire survey in February 2001, for both survey respondents using NGSS and those not using it. (Fig. 3) As for the merits of the system, according to businesses who use NGSS, we can see that it resulted in lessening the burden of those taking the survey, with responses such as "I was able to save time in sending the survey," "It reduced the amount of time need for recording my place of business and its location," "It was labor-saving in that it eliminated work for transcribing data," etc.

Regarding businesses that do not use NGSS and instead make their submissions on paper, as for their reasons for not introducing the system in the present situation, there were many businesses that responded by saying "It seems like it would take more time than doing it on paper." However, because there is a high level of interest in the system, with almost 60% of those companies that do not use NGSS saying they "intend to introduce it in the future," if we continue our educational activities regarding the benefit of the reduced burden for those completing the survey using NGSS, it should be possible to significantly increase the rate of online submissions.

(Fig. 3)

(2) Improvement of System Performance

Using NGSS, from the trial operation stage, various requests for improvement of: the processing capability of examining individual survey responses,   the capability to communicate with all prefectures and those taking the survey, and   operativity have been made, and we are responding sequentially where possible.

First of all, regarding the processing capability of examining individual survey results, because of many claims that the processing speed is too slow, we are currently increasing the number of servers and tackling the improvement of system performance. As for communications capability with all prefectures and those taking the survey, claims have been made that it is difficult to connect to the system or that even when a connection is made, transfer of data takes too much time. Therefore we are carrying out improvement of the Web server performance and improvement of communication lines for dial-up connections.

As well, regarding requests made for improvement of operativity from those taking the survey, the prefectures of Japan, METI's regional branches and its headquarters (Research and Statistics Department), we are conducting maintenance where it is possible to respond, and carrying out sequential release.

In particular, it is necessary to give priority to tackling requests for improvement of operativity and problems with communication lines that serve as the interface with those who take the survey, not only because these are important issues, but also for the sake of increasing the rate of online submissions.

(3) Expansion of Online Information Provision

NGSS is already implementing online provision and circulation of data. However, in the future it will be necessary to aim for expansion of online information provision, and make it possible for all users, including those who take the survey, the prefectures of Japan, METI's regional branches and its headquarters (Research and Statistics Department) to receive survey and system explanatory materials other than data from the home page of the system.

(4) Upgrading Security

As already mentioned, we are taking appropriate security measures regarding operation of NGSS. However, due to the high speed of technological progress, it is necessary to see that the best measures are being taken on a daily basis. Because of the fact that there are businesses that listed security concerns as a reason for not making use of NGSS in the questionnaire survey, it is important to clearly explain our current security measures to businesses.
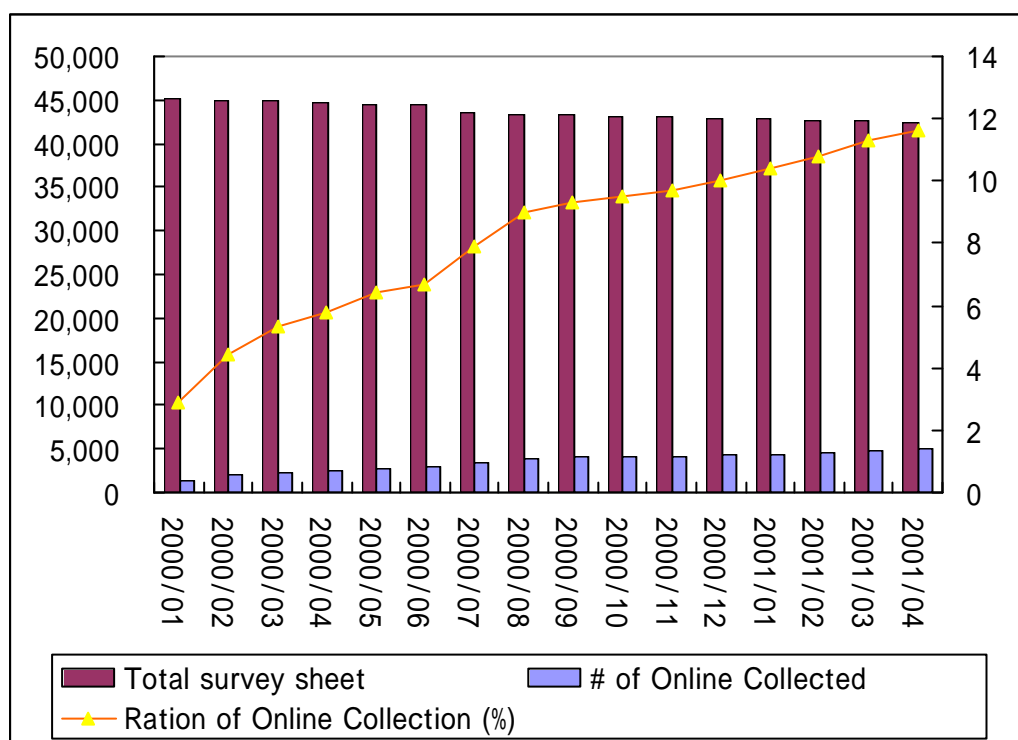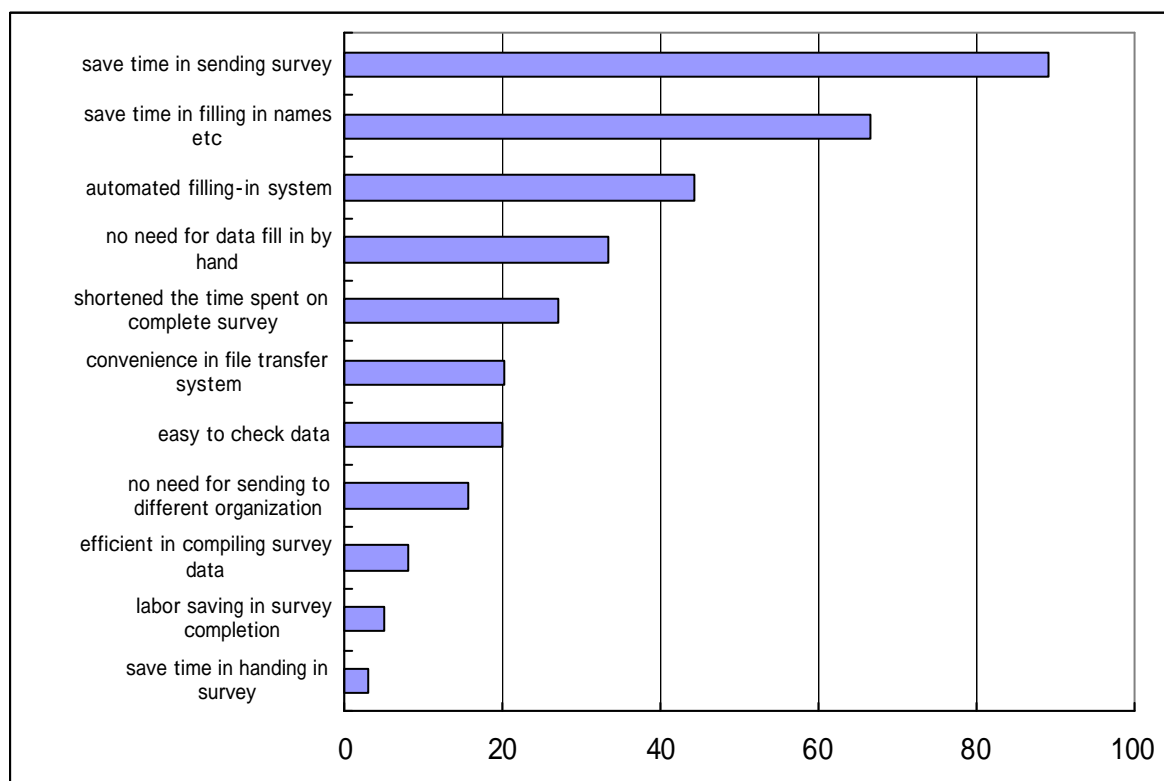
Fig 2: Trends in Online Submission Rate (%)

Fig 3: Merit of NGSS (% of responses in each factor seen as a merit)

Appendix 1. **Statistical Surveys Covered by New Generation Statistical System (NGSS)**
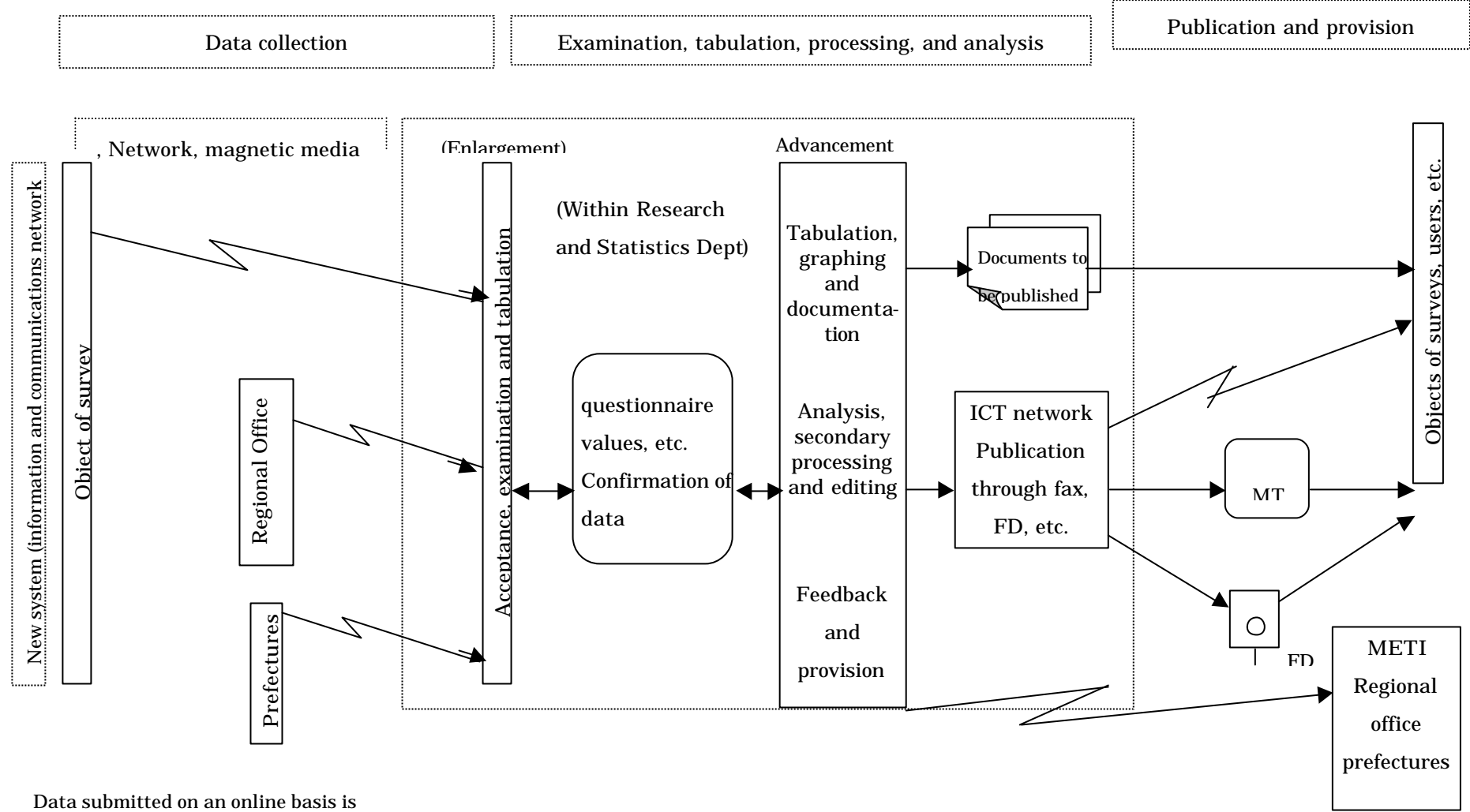
Field	Primary statistics		Secondary statistics
	Census statistics	Current statistics	Indices, analyses, etc.
Mining and manufacturing	Census of Manufactures <annual> Survey of Mining Trends in Japan <annual> Census on Selected Machinery and Equipment, Etc. <once in several years>	*Current Survey of Production* *<monthly>* *Manufacturing Industrial Production Prospect Survey* *<monthly>*	*Indices of Industrial Production* *<monthly>* *Manufacturing Industrial Production Prospects Index* *<monthly>*
		Supply and Demand Distribution Statistics *<monthly>*	
Commerce	Census of Commerce <every five years> (Supplementary survey to be conducted intermediately)	*Current Survey of Commerce* *<monthly>* *(Large-scale wholesalers)* *(Large-scale retail stores)*	*Indices of Commercial Sales* *<monthly>*
			Indices of Tertiary Industry Activity *<monthly>*
Services	Survey of Selected Service Industries <annual>	*Current Survey of Selected Service Industries* *<monthly>*	

Energy	Structural Survey of Energy Consumption <annual>	*Current Survey of Energy Consumption* *<monthly>*	

All industries	Basic Survey of Business Structure and Activity <annual> Basic Survey of Commercial and Industrial Structure and Activity <once in several years>		Input-Output Tables <annual, every 5 years>

(statistics processed by NGSS is shown in *Italics*)

Appendix 2. Diagram of the New Generation Statistics System (NGSS)

Data collection

Examination, tabulation, processing, and analysis

Publication and provision

, Network, magnetic media

(Enlargement)

Advancement

New system (information and communications network

Object of survey

Regional Office

Prefectures

Acceptance, examination and tabulation

(Within Research and Statistics Dept)

questionnaire values, etc. Confirmation of data

Tabulation, graphing and documentation

Analysis, secondary processing and editing

Feedback and provision

Documents to be published

ICT network Publication through fax, FD, etc.

MT

FD

METI Regional office prefectures

Objects of surveys, users, etc.

Data submitted on an online basis is directly entered into a METI computer, and is automatically accepted. METI's regional branches and prefectures confirm the state of acceptance and perform examinations.

Note: MT (magnetic tape) and FD (floppy