

SEMINAIRE

СЕМИНАР

SEMINAR

STATISTICAL COMMISSION AND
ECONOMIC COMMISSION FOR EUROPECONFERENCE OF EUROPEAN
STATISTICIANSDistr.
GENERALCES/SEM.38/28
21 April 1998

ENGLISH ONLY

Seminar on integrated statistical information
systems and related matters (ISIS '98)
(Geneva, Switzerland, 27-29 May 1998)

Topic (iv): automation of data collection and its relation
with the achievement of lower costs and greater efficiency

APPLICATION OF EDI IN DATA CAPTURE FOR BUSINESS STATISTICSSubmitted by the Czech Statistical Office¹**I. ELECTRONIC DATA COLLECTION PROJECT**

1. In order to streamline the process of collecting statistical data on the economy, the Czech Statistical Office (CSO) launched a project on **electronic data collection**. The primary purpose of the project is to enhance the quality of collected data, and to decrease the amount of work and the response burden. It is based on the level of computer technology of the respondents at present (PCs) and advanced means of electronic communication.

2. The basic objectives of the project are as follows:

- Creating a software tool that allows respondents to **complete an electronic questionnaire** using a PC in addition to completing a paper questionnaire;
- **Providing** the software of the electronic questionnaire to respondents **free of charge**;
- Using **electronic mail** and **Internet** technology for the dissemination of electronic questionnaires to respondents to the maximum possible extent;
- Designing an interface to integrate electronic questionnaire into respondent's intra-enterprise information system.

II. WHAT IS AN ELECTRONIC QUESTIONNAIRE

3. An electronic questionnaire (EQ) is a specialized software which creates on a respondent's PC a full view of a particular statistical questionnaire ready to be

¹ Prepared by E. Petrikovits.

filled in with statistical data. The EQ includes the necessary context-oriented explanatory notes and nomenclatures, and detailed instructions on how to use this product. The filled in data can be changed and amended, until the questionnaire is correctly completed. The input data are checked by the same system that is used for data processing at the CSO. This eliminates the need for costly corrections of the questionnaires already collected when they are processed at the CSO. The check sums, which are a standard part of a paper questionnaire, are automatically calculated by the EQ software.

4. The completed and checked questionnaire (on a computer screen) can be sent by the respondent to the CSO in several forms depending on their technological equipment. In its simplest form: (i) the questionnaire can be printed on paper and sent to the CSO by the respondent; (ii) the questionnaire can be copied in a standardized format on a floppy disk and sent by mail to the CSO; (iii) the data file can be sent to the CSO by e-mail (which seems to be the most efficient way).

5. Individual data protection in compliance with the State Statistical Service Act No. 89/1995 is an important aspect of the whole EQ system. A special cryptographic procedure will be included in the EQ software and used for data transfer in electronic form.

6. Respondents can obtain the necessary EQ software free of charge in two ways:

- **on request** - the CSO is ready to send the appropriate software on a floppy disk, or
- **by downloading** via communication lines - either from the CSO BBS System or from CSO web server.

III. PREPARING AND IMPLEMENTING THE PROJECT OF ELECTRONIC QUESTIONNAIRE

7. Up to now, three stages of the project have been completed:

Stage 1 (in 1995/1996) encompassed relatively extensive research among respondents aimed at measuring interest in electronic data collection for economic surveys. The research was based on the available computer technology (equipment with PCS). The respondents' interest was surprisingly high and, in view of our previous experience, unusually positive.

At Stage 2 (1996/1997), the basic required functions of electronic data processing were defined taking into account the high interest of respondents. The crucial decision was that the basic form of communication with respondents remains the statistical questionnaire. At this stage, the software was tailored to one particular statistical questionnaire (P3-04) used for quarterly economic survey in 1997. This solution was used to test the whole system in operation with respondents in terms of technology and organization.

Within Stage 3 (1997/1998) the electronic questionnaire, its preparation and dissemination to respondents, was integrated into the standard technological environment used in the CSO for the preparation of statistical surveys. Twenty-four questionnaires of various periodicity have been made available to respondents for statistical surveys in 1998.

IV. INTEGRATING ELECTRONIC DATA COLLECTION INTO STANDARD TECHNOLOGY

8. With a view to building a unified technological environment for statistical surveys, we were well aware that electronic data processing is not an "independent branch". It must become a module for creating, predominantly automatically, appropriate applications both for respondents and processing in the CSO. This

requirement was based on two factors: (a) the necessity of maintaining full automatic compatibility between data processing software in the CSO and the EQ software of respondents, (b) limited programming capacity.

9. For these reasons, the programme module generating the EQ software package (installation file) of a given questionnaire for respondents was fully integrated into the programme system ProjektMan (product for computer-aided technological preparation of statistical surveys). It means that the programme generating the EQ and the data processing programme in the CSO are identical in terms of required functions. These functions are specified in the technical description of a statistical survey project (i.e., subject-matter and technical specifications in ProjektMan).

10. The module generating the EQ software transforms the ProjektMan specification of a statistical questionnaire into a suitable form to be completed on a PC. This includes the scopes of completed fields, logically conditioned leaps, context prompts on individual items, and the definition of the structure of the output file. The full block of check links is also automatically taken over from ProjektMan, including the distinction between interactive and batch links. The programmer needs only to link the necessary nomenclatures and other files which are not run under ProjektMan in terms of contents. At the same time a manual intervention is possible in the generated parts of software.

11. A unified software ensuring auxiliary functions (installation and deinstallation procedures, installation instructions, printing the questionnaires, prompts, conditions of use, etc.) is attached to the generated EQ. The whole software package is compressed and ready to be provided to respondents. The average size of such a file is about 1.5 MB.

V. PRESENT EXPERIENCE

12. While testing a selected questionnaire in 1997, we gained valuable experience, particularly from respondents' reactions to the software. Judging by the responses received we feel that we had underestimated the elucidatory campaign for EQ use. Data exchanged between the CSO and respondents was mostly on floppy disks. Moreover, a number of problems occurred only in 1998 when we implemented a relatively extensive set of EQ.

13. In the first place, the capacity of the communication lines between the CSO and the Internet provider proved to be insufficient. The second problem encountered was the reliability of Internet with the transfer of files of 1.5 MB. An additional complicating factor is that the CSO addresses a part of respondents every year for the first time. Therefore it is necessary, because sample surveys are used, to repeat the relatively demanding campaign in order to address these respondents and make EQ available.

VI. FURTHER DEVELOPMENT

14. Our experience and the mostly positive respondents' attitude encourage the further improvement of this technology. The EQ technology should be developed in two levels:

- At the **technical level** we plan to complete the EQ software with an accurately defined data interface allowing a direct link to its information system. Its primary purpose is, particularly for questionnaires with financial indicators, to replace manual data input with automation. This function must be part of the respondent's information system; therefore, this function has to be implemented by software companies making these systems commercially.

- At the **organizational level** it is necessary to create a database of respondents (including E-mail addresses) who are able and willing to use the EQ system. This database should then be used to disseminate EQ software directly from the CSO. To a certain extent, this solution is needed because of changes in statistical surveys every year. These changes often require a complete reinstallation of EQ software on the CSO web server.