

**CONFERENCE OF EUROPEAN STATISTICIANS**

**UNECE Work Session on Statistical Data Editing**

(27 – 29 May 2002, Helsinki, Finland)

Topic (iv): Impact of new technologies on statistical data editing

**IMPLEMENTATION OF NEW TECHNOLOGIES IN  
STATISTICAL DATA COLLECTION**

**Contributed paper**

Submitted by the Central Statistical Office of Poland<sup>1</sup>

**Abstract:** This paper describes the Central Statistical Office's (CSO) plans to implement Hand-Held Computers (HHC) with data file transmission capabilities and relevant software to support the gathering of statistical data by interviewers.

**I. INTRODUCTION**

1. To improve the workability of the regional statistical structures in Poland a number of IT investments financed from the national budget have been done in 2000 – 2002 and new investments to be financed in 2002 – 2003 from the Phare funds are in preparation.
2. The plan includes the purchase of HHCs with relevant software and data file transmission capabilities, to support the gathering of questionnaire data by statistical survey interviewers, to improve quality and productivity of data collection.
3. For this purpose a pilot implementation of HHCs will take place at 3 selected Statistical Offices (SO), before such equipping in the next stage all of 835 interviewers over 16 SOs.

**II. NEW TECHNOLOGY IMPLEMENTATION**

**A. Scope of use**

4. It is planned to apply HHCs at first within the most time-consuming statistical surveys conducted by SO interviewers, i.e.:
  - **PRICES** – the objective is to collect every month information on retail prices for calculating price indices. The survey is conducted in 310 areas throughout the country and is based on 2000 representative products and services. There are about 28,000 outlets, where interviewers throughout the country enter yearly about 500,000 prices. The size of the file including data for one area in one year is approximately 5 MB.
  - **ECONOMIC ACTIVITY (BAEL)** – this survey defines the economic activity of the population. The quarterly survey involves approximately 19,000 households. There are two questionnaires: one for collecting data that are household characteristics and one for questions concerning economic activity. The size of the data files is approximately about 2 MB per year.

---

<sup>1</sup> Prepared by Tomasz Pawlak and Maria Wesolowska (m.wesolowska@stat.gov.pl).

## B. Concept of the HHC statistical applications

5. The CSO concept of the a/m HHC statistical applications is based on the following assumptions:
- HHC data communication abilities,
  - equipping each of HHCs by their supplier with a dedicated Work Management System, according to the CSO requirements,
  - statistical survey application programs for the HHC data collection for all interviewers throughout the country will be centrally developed at the Central Statistical Computing Centre Branch at Radom in the dedicated development environment to run under Work Management System and will be distributed via the CSO Intranet.
6. Working with hand-held computers (HHC) interviewers will use data communication with servers connected to the parent SO LAN via dial-up phone cable lines, to receive survey application programs and organizing data files (instructions, addresses, lists of products, etc.) and to transmit collected statistical data files.

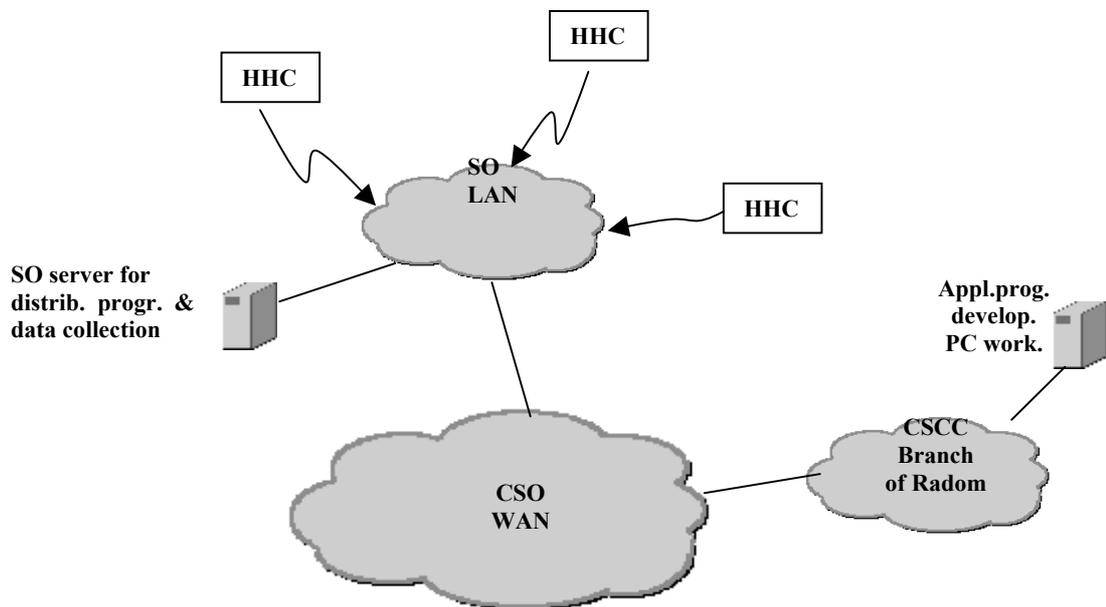


Diagram 1. Planned data communication schema between HHCs and CSO Intranet via parent SO LAN, using dial up phone cable lines.

7. Each HHC will have a dedicated Integrated Work Management System, which will cover management of all universal interviewer operations. This system will include at first management of the data communication via dial-up phone cable lines with the SO server for distribution of the HHC survey application programs and organizing data files and for collection of the HHC recorded statistical data files, as well as management of an automatic installation of the received application programs without intervention of interviewers and further management of starting them, working with them and closing these programs by interviewers executing preparation and real data collection.
8. The HHC integrated work management system will have a simple and comfortable user interface to control all operation functions, whose use will require only minimal training of the interviewers.
9. In the case of data loss or when exchanging a HHC there will be a possibility to download all that is needed to renew a data collection process.
10. The collector will have the possibility of reviewing reports concerning the data collection process and HHC management. The interviewer will have the possibility reviewing monitoring data, but wouldn't have the possibility to change it. Data stored on HHC will be secured. No one except an interviewer that

is HHC owner will have possibility to read it. Received or sent data will be also transferred in a secured way, concealing data files against outside general inspection.

### **C. Data capture applications**

11. Statistical survey application programs will be developed in an external environment to the HHCs, to be automatically installed and further executed under the HHC dedicated Work Management System. Each one will reflect the existing statistical questionnaires therefore enabling entry of data and their automatic validation checks and verification and navigation through the questionnaire itself. The resulting data files will be transferred to the Statistical Office to which the interviewer is assigned.

12. Data will be validated during entering. Errors will be immediately flagged and the data collector will be forced either to correct the data or to provide some explanation.

13. A collector will have the possibility to choose the route, a quotation point (for example sale point) and a representative to enter data. A representative will be also automatically chosen when using a barcode reader. A collector will also have the possibility to define or change the survey organization components such as route, quotation point, representative and so on.

14. There will be the possibility to re-order survey components to appear on screen in the order that is most convenient for the data collector. He or she will have an opportunity to enter comments about survey components too. There will be ability to use predefined codes (for ex. "PO" means comparable in case of representative). He or she will have a possibility to use at any time a context sensitive help.

15. There will be a facility to keep track of errors made by collectors before edits. It could help identify inappropriate instructions, insufficient training and other deficiencies. A collector will have the possibility to review reports concerning efficiency of his work.

### **III. CONCLUSIONS**

16. It is expected that implementing HHC in data collection will improve quality and productivity of the statistical data collection and gives possibility to increase the number of surveys conducted in this way.

17. The following operational advantages of using such HHCs are predicted:

- reduction of errors due to the interactive data validation on the spot,
- reduction of manual tasks compared with filling in paper questionnaires and further typing/keying for data entry,
- avoidance of transcription errors,
- reduction of work time of interviewers.

### **References**

Central Statistics Office of Poland, Phare'2000 project PL0003.14 Regional Statistics - *Hand Held Computer Requirements*