Distr.
GENERAL

ECE/CES/SEM.54/24
8 June 2006

ENGLISH ONLY

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE CONFERENCE OF EUROPEAN STATISTICIANS

EUROPEAN COMMISSION STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT (OECD) STATISTICS DIRECTORATE

Joint UNECE/Eurostat/OECD Seminar on the Management of Statistical Information Systems (MSIS)
Sofia, Bulgaria, 21-23 June 2006

Topic (iv): Review and follow-up to the activities of the Conference of European Statisticians

CAPI SYSTEM AT CENTRAL STATISTICAL BUREAU OF LATVIA

Supporting Paper prepared by Norberts Talers, Central Statistical Bureau of Latvia

1. The system is intended for collection of social statistics data with use of CSB interviewers' section. Until 2006, the data on all surveys collected by interviewers' section was done in PAPI mode. In this mode, the validation procedures were done at data entry stage, which was performed also by interviewers, and in case of errors it was troublesome to correct the value after the interview. Also, the monitoring of interviewers' progress of work was less effective compared to that in CAPI mode. The time period needed to receive the data by CSB was longer in paper mode because of double data entry – on site and later on entering the data in data entry form.

2. The purpose of the CAPI system development is to increase the work quality and timeliness of the interviewer section in a sense that data collection can be made easier and faster for the interviewers.

3. As the data is entered directly into a laptop no more data entry is necessary from paper forms, and the data validation procedures are carried out on – site, while interviewing the respondent.

4. Wireless telecommunications system utilizing the GPRS / EDGE –technology appeared to be the easiest and most inexpensive solution.

5. The GPRS coverage is good in all parts of Latvia.

6. In the current system data to and from interviewer is sent as an FTP-transfer via a GPRS connection. Connection times are few minutes at most and the cost is by the amount of transferred data.

7. The gain of using mobile networks and GPRS is that interviewers are able to send the data virtually any time, for example after receiving the information from every respondent thus making the time from entering the information into the laptop to getting the data to the office very short and diminishing the problem of possible data loss due to some kind of a technical breakdown. Also, the management of interviewers work becomes easier and more straightforward while using CAPI technologies, because of possibility to control the process at any time.

8. At the beginning of 2004 a pilot project on data collection with 5 interviewers using laptops and Blaise was carried out; 100 respondents were questioned on Labour force survey. In 2005, on May, the development of CAPI system at CSB was started. One person was responsible for developing CAPI system, and two programmers were responsible for developing first questionnaire in Blaise. In December 2005 45 laptops were received, 45 interviewers were trained, and from the January 2006 the data in Labour force survey is being collected by using CAPI technologies. From middle of March 2006, EU – SILC survey data collection has been started within the frames of the mentioned system. CSB has plans to move all permanent surveys to CAPI environment by the end of 2007.

9. The system consists of three logical parts (data flow is shown in picture 1):

(a) Case management system on interviewers laptops consisting of two software types;
- for the data entry of questionnaires specialised data entry software Blaise is used;
- for the case management a tool developed by CSB programmers is used;

(b) Data transfer system via GPRS;

(c) Information management at the office.

10. When working with a laptop, an interviewer is working with the case management system with several possibilities:

- Receive the data on new surveys (respondent list and questionnaire itself)
- Send the data about surveyed respondents, and the data itself
- Work with received respondent list, surveying respondents

11. The core case management system is the place where an interviewer can see the respondent list for a survey chosen, with possibilities to set a meeting time for a specific respondent, or to open Blaise questionnaire, to which specific commands are passed to Blaise from case management system and start collecting data. Also, interviewers have possibilities to see the statistics on their work done (how many respondents have responded, how many non responses, or how many not yet questioned). Different filtering functions are prepared to ease the work of an interviewer. Case management system logical structure shown on picture 2.

12. Whenever data sending and receiving is taking place, GPRS is used for that – on each of the laptops GRPS card is installed, which is used for connection to service provider. In case of receiving an update of the respondent list, or receiving completely new questionnaire with

respondent list, interviewers just have to press one button, enter authorization information and wait for acknowledgement information from the system about successful data receival, or about the failure. Technically, a script is activating GPRS connection on a laptop, connecting to an FTP server at office of CSB, and downloading a password protected zip file, from which upon the receival data is extracted. Some file consistency checks are done in order to be sure that the file is not corrupted.

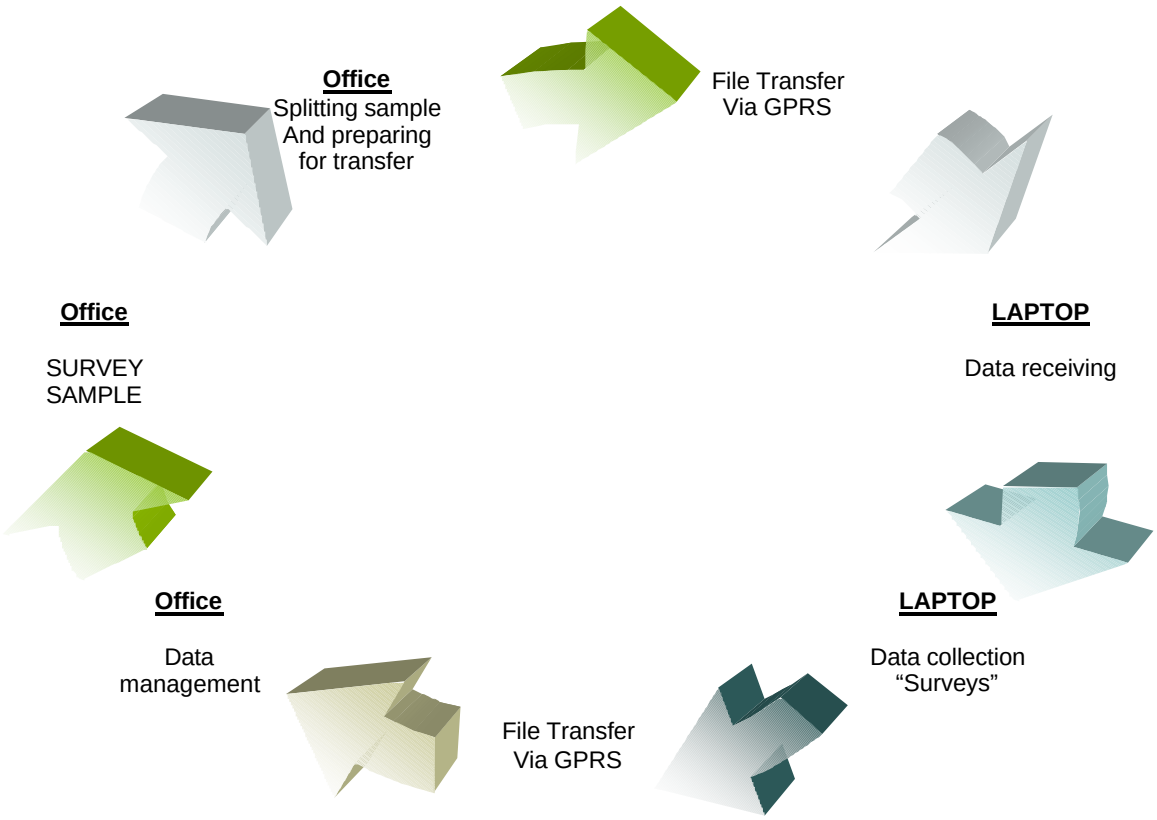
13. When the data is sent from the laptop to our office from point of view of an interviewers it is much the same as in case of receiving the information. Scripts are collecting the needed files to be transferred, zip them in a password-protected archive and put on an FTP server at office of CSB.

14. Information management at the office includes several parts:

- A tool for respondent list separation in portions for each interviewer based on division by territories predefined.
- A tool for archiving of the respondent list part and the questionnaire, and putting that on an FTP server;
- Getting the information sent back by interviewers from FTP server, with possibilities for interviewers' section chiefs see the progress of work on surveying respondents.

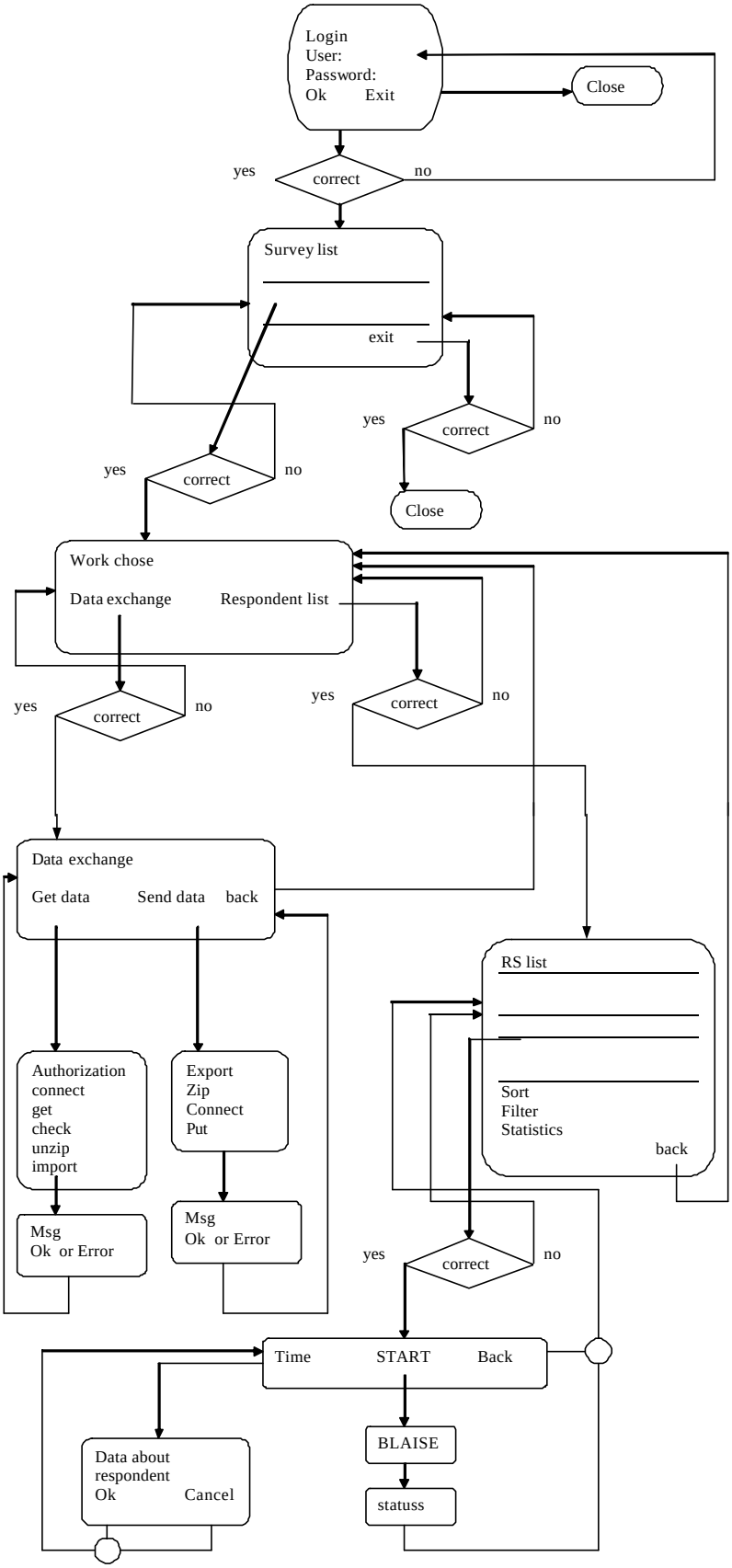
15. There is also possibility to manage the information (e.g. creating additional folders, deleting the information, updating the case management system etc.) that resides on the laptops directly from the office by using scripts that are sent to laptops via the GPRS connection.

16. In near future it is envisaged to implement a simple messaging system that will allow interviewers to receive information from the office and send some messages to the office as well as to other interviewers via GPRS connection. This will avoid using phones in order to contact people, thus diminishing costs.



Picture 1 - Information rotation scheme

17. The picture above shows the information rotation between a laptop and office side of CSB. The essence is that no other means of data transfer than GPRS is used at any time.



Picture 2 - Case management system logical structure