

A PROTOTYPE SYSTEM OF DATA EXCHANGE FOR STATISTICS FINLAND'S CAPI SYSTEM

Vesa Kuusela

Statistics Finland, Helsinki, Finland and

Antti Merisalo

Tele-Mail Software Ltd, Helsinki, Finland

1. Introduction

At Statistics Finland, Computer Aided Data Gathering (CADAG) has been under development for some time. In the first phase, a CATI centre was set up in the spring of 1991. Blaise was selected as the CADAG software. The first study was undertaken in May 1991, and since then five studies have been completed by CATI.

The second phase of the CADAG project was to analyze and to test various possibilities for putting the CAPI organization into operation. At the moment, CAPI activity is still in the planning stage. Statistics Finland has some 160 field interviewers and their central unit is in Helsinki where all of the interviewing is supervised. Only ten of the interviewers have a laptop computer but all interviewers will later be equipped with one.

It became apparent in the very beginning that the exchange of data between interviewers and the central unit was perhaps the most complicated part of a CAPI system. However, it is of major importance and needs to be designed carefully. If the subsystem for data exchange does not function adequately, it both ties up a lot of personnel and is error prone. In the worst case, all the benefits which were behind the decision to establish a CAPI system will be lost.

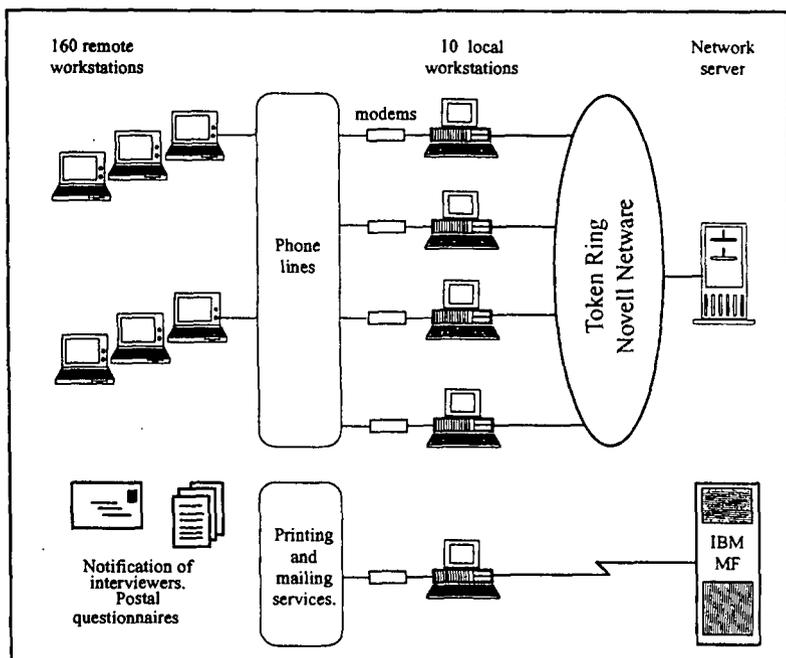
In this paper we describe the prototype of a system designed to enable the interviewers to receive questionnaires and contact information and to return data to the central unit.

2. Hardware

The basic idea behind the system is to use the CATI centre as the centre of services for CAPI as well. At the moment, the CATI centre has ten work stations connected to an i486-based server via Novell Netware and Token Ring. The CATI centre is in use mainly in the afternoon and in the evening. It is seldom used in the morning because many of the people to be interviewed are hard to reach at that time.

Because the work stations are in the CATI centre, there is naturally a telephone line for each of them. The CATI work stations needed only an internal 2400 baud modem to work as service stations. In addition, the modems can be used for automatic dialling, as well, when the feature is added to Blaise. Each work station can be used either in the CAPI service mode (and the modem is programmed to answer incoming telephone calls) or in the CATI interview station mode. This means that there are 1-10 telephone numbers available for field interviewers (see figure 1). The number of service stations is easy to increase when the CATI centre is enlarged. However, the current number of service stations probably will suffice even when all interviewers have laptops.

Figure 1. The hardware configuration of the CATI centre



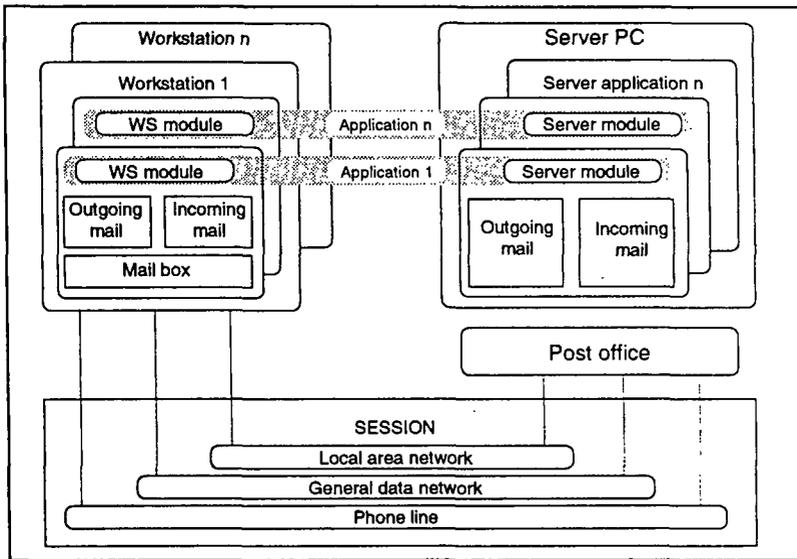
As the CAPI system is still in the planning and testing phase, only ten of the 160 field interviewers have a laptop computer at the moment. The computers are Sharp 4721's with a 20 Mb hard disk and a 1.44 Mb floppy drive. Interviewers also have an external 2400 baud modem. In addition, there are also some parts of Blaise and Turbo Pascal in each machine. The reason for this will be explained later .

3. Software

The software for data exchange was specially designed and made for this purpose by one of the authors (AM). Additionally some public domain software to pack and to transmit data are applied.

One design principle has been that the interviewers contact the central unit when it best suits them within a given time interval. Another principle has been that they need only a few easy commands despite of the fact that the processing capacity of their laptops is used to make the working questionnaire for the study. Only the text (ASCII) form of the questionnaire is sent to interviewers and it is compiled in their machines. This is done in a batch run.

Figure 2. General description of the software system

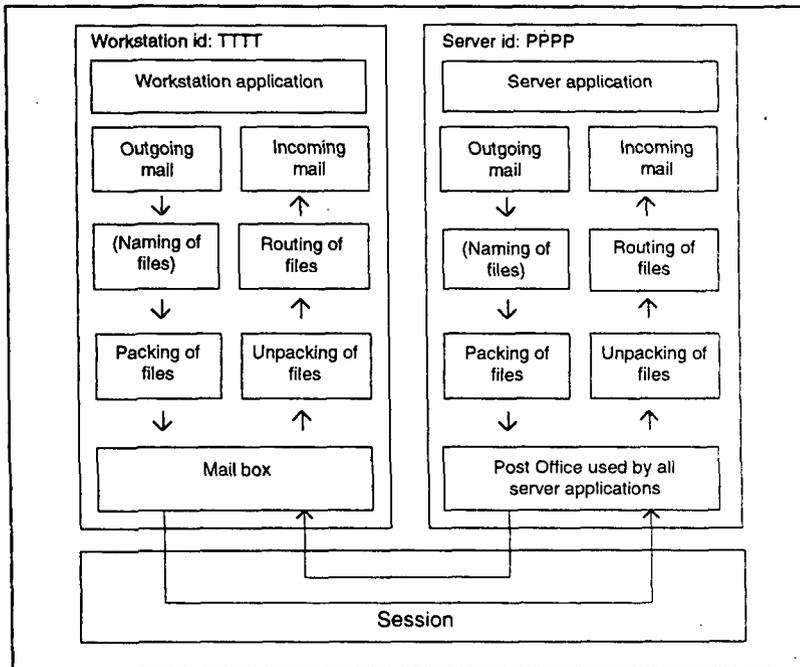


The system is described in figure 2. Every work station may have several CAPI applications at the same time. The server software may be in any machine connected to the LAN of the CATI centre. The machines in the CATI centre may be used both as servers for CAPI and as workstations for CATI.

4. Server Software

The server software is the part of the system which resides in the work stations of the CAPI service centre. The software may be in sending or in receiving mode during the same session. The switching between modes is automatic. A more detailed description of the logic and of the data flow in the system is given in the figure 3.

Figure 3. Detailed description of the data flow in the software system



When in sending mode, the server software prepares questionnaires (text files) to be sent, makes a packet for transmission, furnishes the packet with an address and time-stamp and moves the packet into the box for outbound post in the post office.

When in receiving mode, the server software unpacks the received data and moves them to the directory of the specific interviewer (all interviewers have a directory of their own). This is based on the name of the incoming file. The name is formed by combining the identification code of the interviewer and of the project code of the study.

Additionally, the server software maintains a log of transactions in the system. On some occasions, the server software joins together the separate interviewer data files of a study and moves the files further to be analyzed.

5. Work station software

The work station software resides in the interviewers laptop computers. It receives the packed questionnaire file, unpacks it and checks its integrity, makes a directory for the study and moves the unpacked questionnaire into it and compiles the questionnaire.

When the interviews of a particular study are finished, the interviewer sends the corresponding data files (*.D?? files) back to the CAPI service centre by a single command.

The work station software and the server software communicate with each other and the interviewer does not have to do anything else except invoke the system. Even dialling and the change of the telephone number in case the line was busy is done by the software.

6. Organization of the field work

Contact information (i.e. the respondents' names, addresses, telephone numbers etc.) will be sent to the interviewers by mail. This is the signal for an interviewer to contact the central unit via modem in order to copy the questionnaire from the centre to his/hers laptop. If interviewers have data to send back it will be done automatically at the same session. When an interviewer has received the questionnaire, a batch run will start. The

batch run contains commands to check the questionnaire's syntax and to compile the questionnaire.

When the field work of a study has been completed, the interviewer sends the collected data files back to the central unit. The server has a directory for each interviewer where the software takes the received files.

7. Summary

The CAPI organization at Statistics Finland is partly based on the use of the CATI centre as a service centre for data exchange between interviewers and the central unit. The system is still a prototype with limited testing possibilities but the results of the few tests have been promising. Compared to many other potential solutions, the prototype described above has the following advantages:

1. The volume of data exchanged by modem is kept to a minimum. With 2400 baud modems, the contact times grow rapidly if the amount of transmitted data grows. Long contact times generate large telephone bills, and they are more error prone and inconvenient.
2. The utilization of the CATI centre grows. It is not very rewarding to try to reach people before afternoon. Therefore the first shift usually starts in the afternoon and the CATI centre is not used before that. The morning is a suitable time for data exchange with the field interviewers.
3. The system is inexpensive. The only additional investments are the modems in the CATI centre and the software. Standard modems are inexpensive and may be used in the CATI centre in the future when automatic dialling is added to Blaise.
4. Easy to expand. If the system faces traffic jams it may be easily enlarged by new modems.
5. The system may be later expanded as an office system for interviewers.