

# **New CAI-operation at Statistics Norway**

*Hilde Degerdal and Jan Haslund, Statistics Norway*

## **1. Background**

### **1.1 Our recent CAI operation**

In 1999 Statistics Norway built a new system for Computer Assisted Interviewing out by the field interviewers. The system has been a good operation for our work. However, after a while you find things that have to be improved.

As a system for field interviewers, the main weakness of our system is the communication solution. It was built as a dial-back RAS connection with a reverse proxy server, polled from inside. Firewalls and encryption are also in use. It is a rather complicated solution, but it had to be this way to meet the requirements of the Data Inspectorate and The Protector of Privacy in Norway. As a consequence of the complexity, the rate of unsuccessful connect attempt is too high.

In our system each respondent is sent to an interviewer as a package. If the interviewer for one reason or another is unable to perform the interview, he has to connect again and return it to the central database. Then, in the office, the staff has to find a new interviewer for the respondent and prepare a new transfer. This work is quite time consuming.

Another problem is that when the respondent packages are out with the interviewers we have little control of the collection progress by each interviewer.

### **1.2 CATI**

Another issue is CATI. In 1991 we started a small CATI-operation in Statistics Norway. However, when we provided our field interviewers with laptops in 1995, the central CATI-staff was no longer required. The field interviewers had capacity enough to carry out those interviews as well. It is also convenient to “fill up” the work load for interviewers who have too few hours just conducting face to face interviews.

That was also the situation when we built our recent CAI in 1999. But after a short period of time, in 2000, it was decided to build up a central CATI-staff again. It was a requirement that CATI should be incorporated in the CAI-system in such a way that it should be possible to shuffle respondents between the two systems. What we did then was to consider CATI as one interviewer in the CAI-system. So each day a large amount of respondents are shuffled to and from the “CATI-interviewer”. This is not the best solution, but for a small scale CATI- operation it has worked.

Last year the CATI-operation expanded to about 40 workstations. This meant that it was impossible to have the entire amount of respondents being shuffled within the CAI-system. So today we have two independent systems, and we are not satisfied with that.

## **2. The project**

A project is going on to make a new system. It is called the SIV –project. SIV = System for Interview Activities. The project’s aim is to make a system that meets the requirements described below.

## 2.1 Requirements for the new system

A main issue is that the respondents shall not be distributed out to the laptops. We want them to be stored and kept in one common database. This means that the field interviewers need to work online when they conduct their interviews.

The system must offer possibilities for different levels of dedication of respondents to interviewers. In some surveys, or phases in the data collection, it may be that we want all the interviewers to work in a traditional CATI-manner. In another situation it may be more suitable to have dedicated respondents to interviewers, or groups of interviewers.

Surveys with mixed or multimode data collection design will be more and more frequent. Therefore, we want to build a system that can handle data coming in from different sources (interview, web, paper) in a simultaneous way.

The importance of a stable and safe communication solution cannot be stressed too much. A system as described, with online interviewing, sets high requirements to up-time.

We want to develop a system built up of modules, so that it can be built piece by piece. This will also make it possible just to renew some of the pieces of the system in the future. Another requirement is to have expertise on the whole system in-house. The system must be built with focus on user-friendliness, and the whole process shall be documented in a user-friendly manner throughout.

## 2.2 Organization of the project

There are a lot of issues to deal with in the project. We have divided the tasks into five subprojects.

1. **Matters related to the interviewers.** This subproject shall develop strategies on levels of dedication of respondents to interviewers. They are going to decide on matters concerning the interviewers working conditions, such as a payment system. Another issue for the group is to evaluate software, hardware and access that can be useful for the interviewers in their work. The last main task for this group is to make good guides for the interviewers on registration of a contact with a respondent. It is important that all interviewers follow the same rules for registration since they are going to handle the same respondents.
2. **Workflow in the survey projects.** We want more standard methods and more streamlined workflow in our work. This subproject shall describe routines; make templates and guides that will form a framework for that. They will give input to the developers on data that are needed to support the data collection process.
3. **Development of the core of SIV.** This group is going to design and develop the central database and operation in the office, the interfaces both in the office and for the interviewers. Their work is going to be done on the basis of requirements from subprojects 1 and 2.
4. **Communication system.** This subproject shall decide and develop solutions for communication, telephony and hardware.

5. **Training of the users.** This subproject will develop a training module for the interviewers and for the staff in the office on the new system. They are also responsible for the quality of the documentation of the entire system.

In addition to the subproject groups we have a group of interviewers and a group from the central staff as reference groups. The steering group is the Department of IT and Data Collection management team, supplemented by two representatives from the labour unions.

### **2.3 Schedule**

We aim to have a new operation in production by 1 April 2008.

## **3. Ideas so far (July 2007)**

### **3.1 Issues to consider**

An important part of the project is of course to make a stable communication solution of good quality. We plan to equip the field interviewers with broadband. In order to reduce the costs, we are also considering the use of IP-phones. That also means that the lines have to be of good quality.

The work on designing the solution has started, there are some ideas, but the main work on this will be done during late autumn or winter.

An issue we have focused on is the system for face to face interview. As mentioned, we want an online solution with the respondent kept in the central database. However, we have done some investigation and concluded that the coverage on wireless connection throughout the whole country is not good enough. So we can't make a solution that requires online interviewing also for face to face interviews. Therefore we will need to build functionality for downloading questionnaires and respondents to the interviewers' laptops, and subsequent synchronizing of the collected data back in the database when the interview has been conducted. We want to have control of all the respondents' status all the time, so we want the appointments to be made online, also for the face to face interviews. That will also make it easier for us to ask another interviewer to do the interview if the first interviewer is unable to do it.

The change from individual respondents to a shared list of respondents for telephone interviews will be the most noticeable change for the interviewers. This will make some challenges. The most important one, as for all CATI-organizations, is how to match the capacity of interview work hours to the actual workload. We think however, that our challenge will be even worse, since the interviewers are not in-house. We will obviously need a monitoring tool for the supervisors. The supervisor's work will require a possibility to direct the interviewers into a survey that needs them. Our idea is to make some parameters for the system that will offer a ranked list of surveys, which is presented to the interviewers when they log on to the system. The list should be read like this: green surveys means "go on", yellow will be "if you want, you may choose this one", red ones have already too many interviewers logged on to them, and will not be possible to start. In addition the supervisor may need a possibility to send "blips" out to the interviewers to advise them to switch to another survey, when that is needed.

Another worry is whether the interviewers' sense of responsibility for making good appointments and persuasive arguments will be the same, when they have common

respondents, compared to the situation of today, where they have their own respondents. Some of the field interviewers have big concerns about this issue. Our experience for telephone interviews in present operations is that CATI does not have a significantly lower response rate than field interviews.

### 3.2 Screens

For the **interviewers** we first want a screen that offers:

- A list of CATI-surveys with common respondent lists. The colour of the printed name of the survey shall indicate which of them they should select.
- A list of their own appointments for today
- Buttons to:
  - Start a selected CATI -questionnaire
  - Transfer respondents for face to face interview
  - Synchronize data from a conducted face to face interview
  - Check if more respondents for face to face interview are available in the neighbourhood. This may be useful if one interviewer has idle capacity, and another interviewer in the area is absent.
  - Open a list of their respondents for face to face or telephone
  - Open a list of their appointments. Own appointments are mainly appointments with their respondents for face to face interviews. Perhaps it should also be possible to transfer CATI-respondents to this list. This functionality may be useful e.g. if the respondent wants to be interviewed at an odd time, which is ok by the specific interviewer.
  - Look up a respondent in the database by name, id, or phone number. It should also be possible to start the interview with the actual respondent.
  - Reports. Some reports will be available, that will give the opportunity to compare own results, and time use, with the average ones.
  - Report work hours
  - Report travel log
  - Questionnaire testing. It would be useful to use interviewers with idle capacity from the entire interviewer staff to test questionnaires. Today a limited team, living near the office, mostly does this job.
  - Read survey manuals. Maybe the survey can't be selected from the list by the interviewer before he has read the manual.
  - Open a questionnaire for training. Maybe the survey can't be selected from the list by the interviewer before he has trained on the questionnaire.

- Access to available web questionnaires. The interviewers should know which surveys have a web questionnaire available, so they can offer the respondent the option to fill this in. It may also be useful for the interviewer to access a certain web questionnaire to help a respondent filling in a questionnaire.

In the interview interface, when calling the respondent the interviewer shall have access to as much information about the respondent and earlier contact (-attempts) as possible.

One of the first things the interviewer has to ask the respondent about is if he has read the letter to respondents. If he hasn't the interviewer can read it for him, or send him a new letter. It must be possible from the system to send a new letter as an attached pdf to an email, or just to click a button, and a message will be sent to the office, telling which letter has to be send to whom for sending paper letters.

For the **supervisors**, we want a screen with a list of ongoing surveys with number of available respondents in daybatch, appointments, and number of interviewers logged on to it. A window within the screen will continuously show messages coming in from the interviewers.

The screen must also have buttons to:

- Get an overview over appointments spread in time
- Give higher/lower priority to a survey if it has too few, or too many interviewers logged on to it
- Change CATI parameters for a survey
- Inspect daybatch for a survey
- Make new daybatch for a survey
- Send a message to an interviewer or a group of interviewers, e.g. to switch survey
- Stop interviewing on a survey
- Look up a respondent in the database by name, id, or phone number. This may e.g. be used to change appointments, or inspect the data
- Do tracking of a respondent. Searching in population register or telephone register to find updated information about the respondent
- Look at a surveys questionnaire. To advise an interviewer or to be able to persuade an interviewer what they need to know about the survey
- Read survey's manual. Same as above
- Access to available web questionnaires

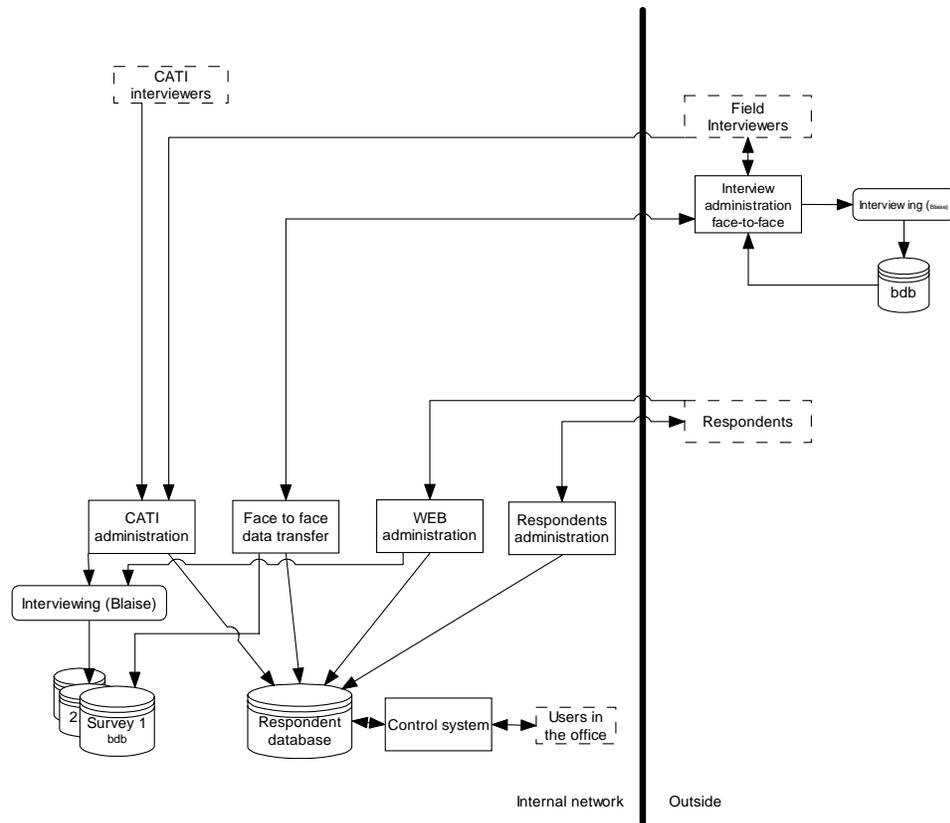
We have also some ideas for a respondent's page. The url of such a web page should be given in the letter to respondents. The page shall offer information about Statistics

Norway, about the survey and about privacy protection. It shall be possible for the respondent to give additional information e.g. telephone number on which the interviewer can reach him, and perhaps also time of day or week he is easiest to come in contact with. We want this information to update the respondent information in the database. Perhaps the web page after the data collection period can give some of the main results from the survey.

When the interviewers are out on face to face interviews they have to fill in a travel log in order to get money back for their travelling expenses. It's mostly driving expenses, so we have started to investigate a solution with GPS and logging, instead of manually filling in a form for travelling logs.

### 3.3 Sketch

As we see it now an overview over the data collection system will be something like this



### 3.4 More ideas

Hopefully we will have some more concrete ideas that we can present in September. The intention to present it in the conference is to get some feed back on our outlines, and perhaps get some additional ideas.