

Business Trends  
Transport and Tourism  
Mr. Ville Vertanen  
ville.vertanen@stat.fi

2 September 2008

Session 4: New approaches to data collection from supply side

## ***Automated Data Collection in Accommodation Statistics***

*New modes of responding that are less burdensome for the data suppliers have been developed to ease the response burden and speed up data collection. Collecting data has become easier as a result of the introduction of an automated reporting system as an alternative to a questionnaire. This new approach also allows us to increase international co-operation and comparability. Sweden, Norway, Denmark, Iceland and Finland have launched a co-operation project which aims to create a common Nordic Automated Data Collection system for Accommodation Statistics.*

Data on guest arrivals and overnight stays accumulate daily in the information systems of accommodation establishments. Hotel systems store the information from each individual reservation. The stored information includes, for example, date of arrival, date of departure, nationality, number of guests, room charge, etc. The data stored into the hotel systems include all the information that is necessary for Accommodation Statistics.

The aim of the Nordic group is to define a Nordic automated reporting system which compiles the Accommodations Statistics monthly report automatically from the hotel management information system and then sends it as an encrypted, electronic report to the National Statistics Office. The report created will be the same in every country, thus minimising the amount of work to the software suppliers, hotel chains or whoever is implementing this system.

According to the aim, one Nordic file, the Nordic group decided to collect the data at the booking level. This ensures that all countries can aggregate the data as they need it for their purposes. The data file also includes some information that only some of the countries need; the other countries can simply ignore this information.

### ***Definition of the automated data collection***

In the automated data collection, statistical information is generated from the informant's management information system (e.g. hotel software) into a specified XML file. When the data are in XML format the file is sent as an encrypted electronic transmission to the National Statistics Office. The respondent is informed if the transfer was successful or if it failed. The informant is able to check the data and revise information, if necessary. Automated data collection is thus not fully automatic as it leaves the informant in control.

In the Nordic model the content of the XML file generated is defined by the Nordic countries and its format is determined by the XML Schema. The idea is that the file is identical for all the participating countries. Consequently the hotel chains or software suppliers only need one procedure that works in every country. The file is intended to be a compilation of the questionnaires from the Nordic countries.

Business Trends  
Transport and Tourism  
Mr. Ville Vertanen  
ville.vertanen@stat.fi

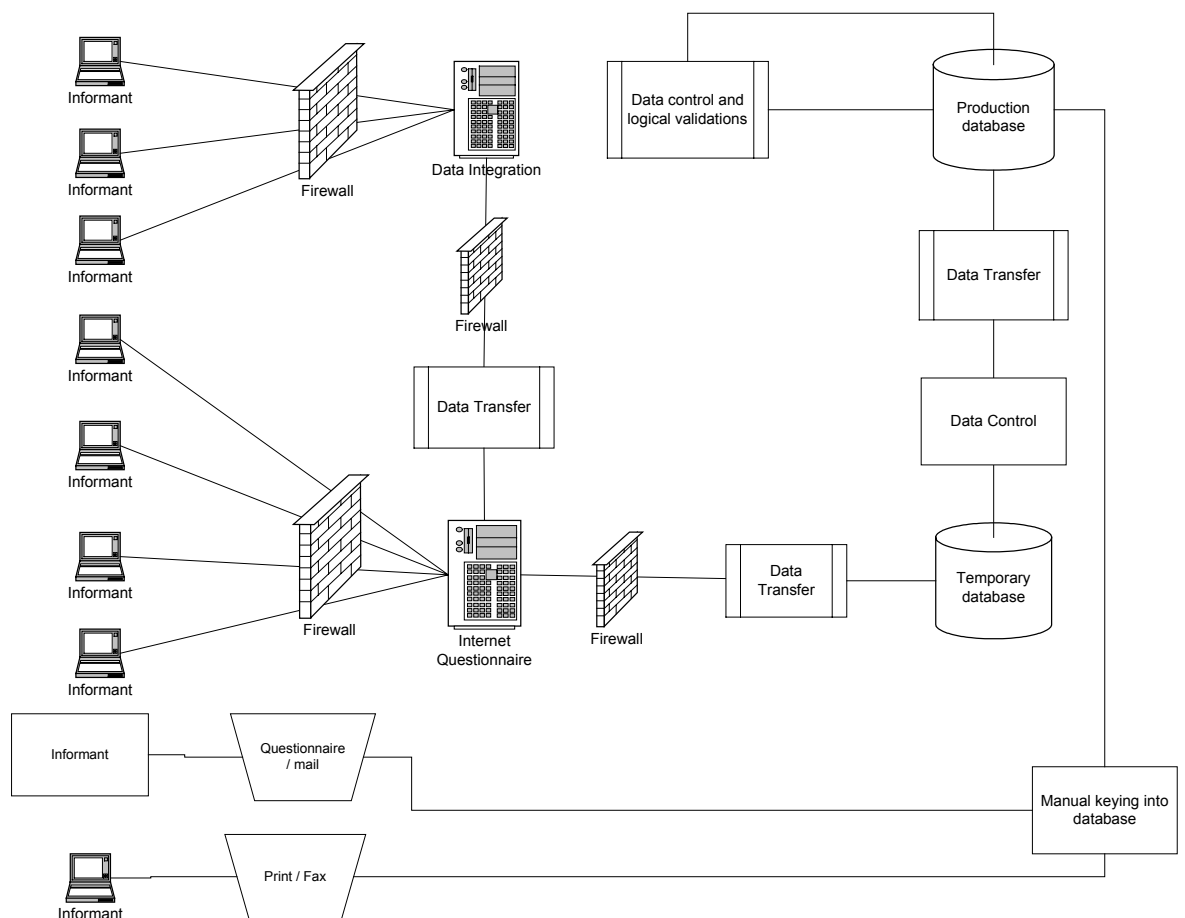
2 September 2008

The technology used in automated reporting also facilitates the sending of compiled reports to other destinations chosen by the accommodation establishment, such as the central, regional or administrative office of its own chain. On the other hand, together with the software suppliers the accommodation establishments can produce tailored reports for the precise needs of their own hotel chain. In addition, their sending can be incorporated into the same automation that sends the data to Statistics Finland. Thus, accommodation chains can exploit the same, ready defined function to collect data on the accommodation activities of their own hotels at short notice. Because the reporting happens direct from the user interface, the data can be reported immediately once the month has changed.

*The data collection process of electronic data transmission, Finnish architecture*

The electronic data collection for Accommodation Statistics comprises two collection alternatives. One is the Internet-based questionnaire and the other the automated data collection. Figure 1 describes the electronic data collection process. This process differs slightly in each Nordic country. The process presented here is the one used by Statistics Finland.

Figure 1 Diagram of the electronic data transmission



In Figure 1, the Data Integration server represents the server for the automated data collection and the Internet Questionnaire is the server for the

Business Trends  
Transport and Tourism  
Mr. Ville Vertanen  
ville.vertanen@stat.fi

2 September 2008

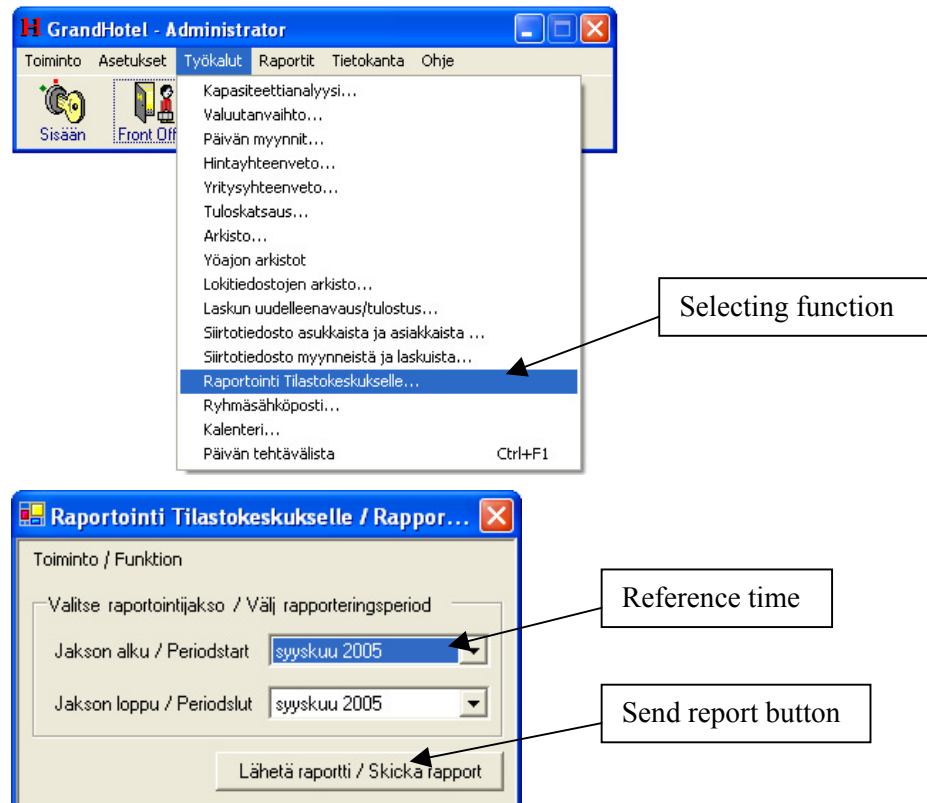
Internet-based questionnaire. If the respondent replies using automated reporting, the received data are transferred to the Internet-based questionnaire within a short delay. Thus the data suppliers can use the Internet-based questionnaire to view the data they have sent.

The team responsible for the Accommodation Statistics can transfer the data from the Internet questionnaire to the production database when required. Before they are saved in the production database, they are subjected to a number of verifications of logicity. For this purpose we have also introduced an application for examining erroneous or significantly deviating data. If critical errors are found the data are verified with the data supplier and then corrected.

### User value and the benefit for authorities

The first version of automated data collection at Statistics Finland was launched in 2005. The positive feedback we received gave us courage to start the more ambitious Nordic co-operation which also implies more users for automated data collection. Respondents have been pleased that they no longer have to fill in questionnaires and the answering is made easy. Figure 2 illustrates the answering process in one software. To answer to the Accommodation Statistics questionnaire, all that is required of the respondents is that they initiate the transfer by selecting the time period and pressing the send button, as demonstrated in Figure 2.

Figure 2 Example of Automated data reporting with Hotellinx software

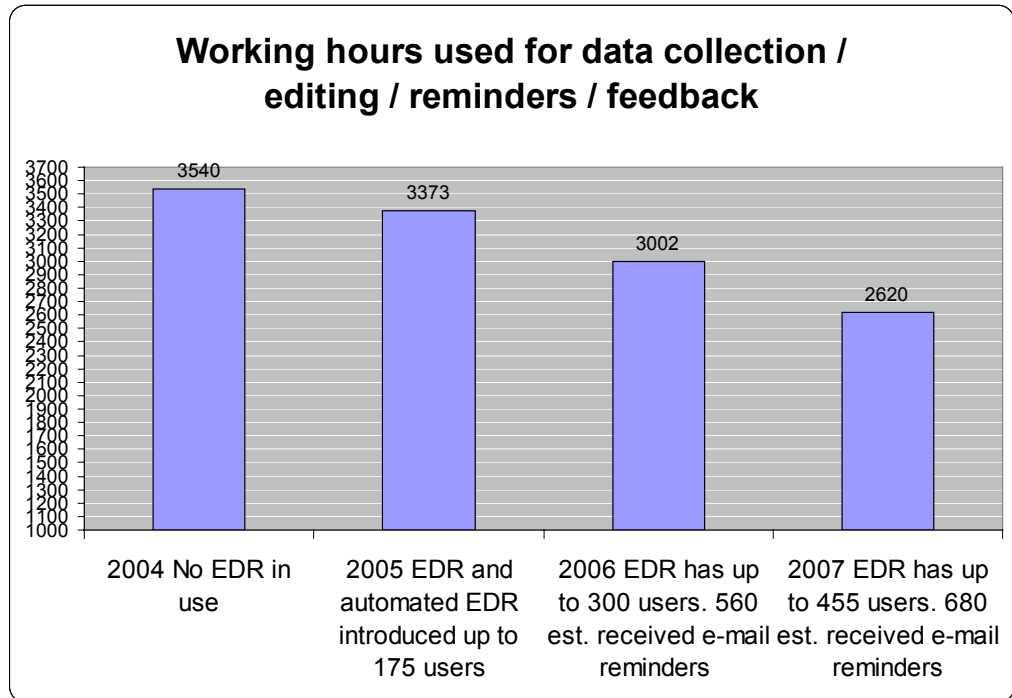


Business Trends  
Transport and Tourism  
Mr. Ville Vertanen  
ville.vertanen@stat.fi

2 September 2008

Automating the data collection eases the response burden tremendously, but it also reduces the burden at the National Statistics Office. As the responses are received in electronic form, no manual input is needed in saving the data into the database. The data must be checked with various routines, but automated data collection does reduce manual work also at statistics office. Figure 3 reports on the reduction of working hours since electronic data reporting has been introduced at Statistics Finland. The reported numbers include also responses coming via the Internet questionnaire. At the end of 2007 a total of 13% of all received answers came through the automated data reporting system and 45% via the Internet questionnaire.

Figure 3. Working hours used for data collection, editing, reminders and feedback (Finland)



\* EDR = Electronic Data Reporting

Since the introduction of the Internet questionnaire and automated data reporting, the number of working hours has fallen by 26% and continues to fall further. For the National Statistics Office it does not matter if the data arrive via the Internet questionnaire or the automated data reporting system. The work involved remains quite the same for the National Statistics Office, but the difference for data providers is significant.

**Conclusions**

As discussed in the previous chapter, Electronic Data Reporting reduces work at the National Statistics Offices, but if it is done with an automated data collection system it also reduces the response burden. Since their introduction hotel softwares have been able to print out paper reports. In the automated data reporting system these same reports are printed into an XML file and then sent to the server defined by the relevant National Statistics Office.

Business Trends  
Transport and Tourism  
Mr. Ville Vertanen  
ville.vertanen@stat.fi

2 September 2008

Because many of the software suppliers operating in Scandinavia are international software houses, the automated reporting described here could be introduced also elsewhere. For example, Scandic Hotels are piloting the Nordic model of automated data reporting. Scandic also operates in Belgium, the Netherlands, Germany, Estonia and Lithuania. Thus the system they have built could be piloted also in these countries. The feedback we have received from the hotels has been positive. Implementing this kind of a system is also a benefit for hotel chains like Scandic. Instead of sending reports from every single hotel, Scandic can in future send the reports centrally from their head agency.

From the perspective of software suppliers and further diffusion of automated data collecting it would be a major advantage if transmissions of the same format could be received in as many countries as possible. In this way automated data collection could be introduced as rapidly as possible into as many programs as possible. If each county defines its own data transmission rules, the software suppliers have to produce a separate transmission program for every country.

If the reporting system is identical in different countries, then the data also come from exactly the same process. This would imply that the data and definitions in these countries are also fully comparable, at least as regards the automated data collection.

We could create international standards for data transfer and file description. Software vendors would then need to build only one system which would work in all countries. Consequently, international implementation would be much faster than if we were to have different definitions in all the countries. There would only be one standard file format which would include all necessary data. The idea would be the same as in the Nordic model: Take what you need and leave the rest.