

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

**EUROPEAN COMMISSION  
STATISTICAL OFFICE OF THE  
EUROPEAN COMMUNITIES (EUROSTAT)**

**Joint UNECE/Eurostat work session on statistical data confidentiality**  
(Geneva, Switzerland, 9-11 November 2005)

Topic (i): Web/on-line remote access

## **MONA – MICRODATA ON-LINE ACCESS AT STATISTICS SWEDEN**

### **Invited paper**

Submitted by Statistics Sweden<sup>1</sup>

**Keywords.** Microdata, MONA, legislation

### **1. INTRODUCTION**

Increased use of microdata requires improved possibilities of providing better data to meet the needs of users. It is vital for National Statistical Institutes (NSIs) to assure that the wealth of stored microdata can be fully utilised by researchers and other authorised users. By and large, the access to microdata means that investments made in official statistics give higher return.

Technological advances in hardware, software, data documentation and the Internet have already largely increased the possibilities to improve the access to microdata, but new possibilities appear every day. However, as the issue of confidentiality protection goes hand in hand with use of microdata, a balance is needed between use of microdata and confidentiality.

### **2. LEGISLATION**

Data confidentiality is guided by two major aspects which both are necessary requirements in order to meet the requests from researchers:

(1) general rules (guidelines, screening procedures, contracts, regulations and laws, etc.), and

(2) technical and practical measures for the same purpose.

The legislation concerning confidentiality and protection of individual's integrity is of importance for the possibility for the NSI to provide access to micro-data. The legislation provides the limits for release of data for e.g. research purposes and underpins and constitutes administrative and technical safeguards for legal founding. Specific legislation of importance is the Statistics Act and the Data Protection Acts. To this specific legislation, the current EU legislation with respect to statistical confidentiality should also be added.

#### **2.1 EU legislation**

The Council regulation (EC) No 322/97 of 17 February 1997 on Community Statistics contains rules that are important for the use of information collected for community statistics. According to the regulation data used by the national authorities and the Community authority for the production of Community statistics shall be considered confidential when they allow statistical units to be identified, either directly

---

<sup>1</sup> Prepared by Lars-Johan Söderberg

or indirectly, thereby disclosing individual information. To determine whether a statistical unit is identifiable, account shall be taken of all the means that might reasonably be used by a third party to identify the said statistical unit. Confidential data obtained exclusively for the production of Community statistics shall be used exclusively for statistical purposes unless the respondents have unambiguously given their consent to the use for any other purposes. However it is possible to allow access for scientific purposes to confidential data obtained for Community statistics.

Of importance for the processing, including release of data, is also the Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (the Data Protection Directive). The object of the Directive is to strengthen data protection, e.g. the legal protection of individuals with regard to automatic processing of personal information relating to them. The Directive has been implemented in all the Nordic countries. The Directive applies to computerised personal data and personal data held in structured manual files. It applies to anything at all done to personal data processing. The new term, "processing", covers all types of processing of personal data, including registration, storing, disclosure, merging, changes, deletion, etc. According to the Directive data must be: – Processed fairly and lawfully. – Collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. However, further processing of data for historical, statistical or scientific purposes is not considered as incompatible.

- Adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed.
- Accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that data which are inaccurate or incomplete, with regard to the purposes for which they were collected or for which they are further processed, are erased or rectified.
- Kept in a form, which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed. Personal data can be stored for longer periods for historical, statistical or scientific use.

## **2.2 The legislation in the Nordic countries**

The protection measures applied to confidential data obtained for statistical purposes are based on several legal acts and directives. However, it should be noted that access to statistical micro data for research or other purposes is a part of NSIs duty service and is not an obligation given by the law. In the Nordic countries there are specific and modern Statistics Acts regulating the use of statistical information. Icelandic law with regard to statistical information have not been updated since 1913. However, no specific legal acts deal with Statistics Iceland in cases of access to micro data. In that respect official statistics in Iceland take account of the general acts on data protection and more recently on acts for the protection of individual's confidentiality. In general, given the lack of rules concerning access to micro data, Statistics Iceland is preparing guidelines in order to meet requests for micro data by external users. Statistics Iceland takes notice of two main aspects in cases of micro data: (1) Common rules internationally on good practices for handling of micro data, and (2) specific rules and practices of Statistics Iceland thus far.

Data collected for statistical purposes, in accordance with any prescribed obligation to provide information, or which is given voluntarily, may in principle only be used for the production of statistics. There are exceptions that enable access to data for research purposes and public planning. However, a condition for the use for research is that there is no incompatibility between the purpose of such processing and the purpose for which the data was collected. The processing of data, which includes release of data, must also be in accordance with the regulation concerning protection of individual's integrity.

Besides the Statistics Acts there are specific Personal Data Acts<sup>2</sup> that apply to the production of statistics and the release of micro data. The Acts are based on the Data Protection Directive and contain rules about

the fundamental requirements concerning the processing of personal data. These demands include, inter alia, that personal data may only be processed for specific, explicitly stated and justified purposes. Very stringent rules apply to the processing of sensitive personal data. Sensitive personal data may be processed for research and statistics purposes, provided the processing is necessary and provided the public interest in the project manifestly exceeds the risks of improper violation of personal integrity. Furthermore in Denmark, Norway and Sweden processing of sensitive data for research purposes needs approval. A scientific project involving processing of sensitive personal data is in these countries subject to notification to and approval by the Data Inspection Agency before such processing can commence. This applies to all surveys, whether conducted by a public administration, individuals or enterprises. (In Sweden the approval of the National Data Inspection Agency is not necessary if a research committee has approved the processing.) If the Data Inspection Agency approves the processing, personal data may be provided to be used in research projects unless otherwise provided by the rules on confidentiality. This means that the NSI may take other issues into consideration even if the Data Inspection Agency (or research committee) has approved the processing of data. Data obtained for statistical purpose are declared as confidential, when they allow statistical units to be identified, directly or indirectly and thereby disclosing individual data. Also anonymous data can be confidential. Statistical data are confidential irrespective of source. Also, data taken from public administrative sources are confidential while in the possession of the NSI. The confidentiality rules are the same irrespective of whether data concerns individuals or enterprises. Under the main rules, access may be granted in forms which do not allow direct or indirect identification of people or other data subjects such as enterprises. However, confidential data may be released to a third party for the purpose of statistical surveys and scientific research. In Finland it is not generally possible to provide access to data when units can be disclosed directly or indirectly.

According to the legislation in Denmark, Iceland, Norway and Sweden, statistical data may even be released with identification data for these purposes. In Finland personal data on a person's age, sex, occupation and education may in exceptional cases be released with identification data for research purposes. One condition in all countries is that access to confidential data for statistical or research purposes must not cause any damage or be detrimental to the data subjects. In practice this means that the NSIs only provide access to anonymous data or de-identified data.

The Nordic countries also have special public business registers that contain some common primary information about enterprises. These registers are (except in Denmark) administered by the NSIs and can also be used for other purposes than statistics or research.

When data has been collected in a voluntary survey the respondents in the statistical surveys must give consent to the release of the data. It is the NSI that decides whether data may be released for research purposes. However in Norway access for other purposes than statistical must be approved by the Data Inspection Agency. The Agency has given general permission to Statistics Norway to provide access to micro data for research purposes and for public planning. The Data Inspection Agency may nevertheless make exceptions to such obligation of confidentiality for certain types of information if they find it in conflict with the Data Protection Act.

The obligation of confidentiality will also – according to the law or by imposition of a duty of non-disclosure – apply to the recipient of the data. The NSI may also impose a restriction limiting the researchers right to re-communicate or use the information. Breach of confidentiality restrictions is punishable by simple detention or imprisonment. In Sweden, however, it is not possible to impose restrictions when data are released to another authority. It is therefore important for Statistics Sweden to take into consideration if the data will be confidential according to the Secrecy Act also at the authority receiving data. If not, any one who so desires can have access to the data because of the authority's obligation under Chapter 2 of the Freedom of the Press Act to provide personal data that are not confidential. However, there are rules providing that confidentiality accompanies data to another authority in special situations e.g. if an authority, for research purpose, receives information from another authority where the data is confidential, the confidentiality will apply also within the receiving authority. However, there are no such rules concerning release of data for statistical purposes or public planning.

In Finland a new Act and Decree on the Openness of the Government Activities came into force in 1999. This legislation contains comprehensive provisions on good practice on information management. For instance the Decree includes a detailed list of general data protection measures for confidential data. Statistics Finland like all the other government authorities has to implement these measures by the end of the year 2004.

### **3. REGISTERS AND MICRODATA IN THE NORDIC COUNTRIES**

The Nordic countries have a long tradition of collecting administrative data and transforming these data to registers suitable for statistical use. The production system and the statistical information system in the Nordic countries are to a great extent based on a number of large administrative registers. However, much is needed to transform administrative registers into high quality statistical registers. In addition, the register system also includes a number of survey-based registers, known as final observation registers (e.g. results from the Labour Force Surveys).

Microdata suited for researchers must be standardised and of high quality. The Nordic countries have compiled a number of integrated registers based on several registers and suitable for analyses and research purposes. The longitudinal integrated register "Louise" from Sweden containing anonymised microdata on individuals and families regarding their education, income and employment might serve as an example. This register includes annual data on all adults in Sweden from 1990 and is updated each year. Such an integrated database offers rich possibilities to carry out different analyses. In the future we see that via the Statistical Data Warehouse we can offer these types of integrated registers "on demand".

Over a number of years some of the Nordic NSIs (Norway, Sweden and Finland) have distributed anonymous microdata to a large number of research institutions and authorities using magnetic tapes, CD-Rom discs, DVD discs or other formats. The volume has increased at the same time as the number of releases/assignments has increased. Denmark has in the past only allowed access to microdata on-site at Statistics Denmark.

### **4. CONFIDENTIALITY**

Confidentiality protection of individual and business data is one of the main principles in official statistics and must be addressed when discussing microdata. The individual is entitled to be protected from unacceptable intrusion into personal privacy. At the same time the individual's need for protection must be balanced against legitimate needs for using information connected to society, such as for statistics and research. The legislation concerning confidentiality and protection of privacy of individuals is of importance for the possibility for the Nordic NSIs to provide access to microdata.

The use of statistical information is normally regulated in legislation and/or in a code of practice. In the Nordic countries there are specific legislations regulating the use of statistical information. According to these legislations, as a main principle data collected for statistical purposes, may only be used for the production of statistics. In addition, access can also be provided for research purposes and public planning. The processing of data, which includes release of data, must also be in accordance with the national regulation concerning protection of the individual's privacy and with the current EU legislation with respect to statistical confidentiality.

All data, including anonymous data, obtained for statistical purposes are confidential. Furthermore, statistical data are confidential irrespective of the source. According to the legislation in Sweden and in other Nordic countries, it is prohibited to disclose confidential data to unauthorised users. According to the main principle, confidential data may be released to a third party only for the purpose of statistical surveys and research. Access may only be granted in forms that do not allow direct or indirect identification of individuals or of other data subjects such as enterprises. In practice, the Nordic NSIs only provide access to anonymous data or microdata without name, address and identification number.

Regarding the use of microdata, legislation in the Nordic countries does not contain any specific rules that restrict the way of releasing microdata. As long as the general requirements in the legislation are fulfilled, the most suitable method can be chosen.

## **5. RECENT DEVELOPMENTS**

Some years ago several of the Nordic countries decided to improve access to microdata. A basic goal was to have a functional and secure way of providing microdata from Denmark and Sweden. Furthermore, such a system should be capable of handling large data sources securely for both the NSIs in Denmark and Sweden and the research community.

In 2004 Statistics Sweden formed a new organisational unit called Register coordination and Microdata access at the Department of Research and Development. Furthermore, a development project was started to investigate whether a new technology for remote access to microdata using Server-Based Computing would be feasible at Statistics Sweden, and if it would be in accordance with Swedish law. A close cooperation was established with some representatives of the Swedish research community to find out their needs and objectives regarding access to microdata. These contacts confirmed the need for metadata as well as the importance of testing the security solution at both ends. In addition, the project internally investigated the number of statistical products that could be handled by the new distribution method.

The results of this development project were very positive and since 2005 Statistics Sweden has a new system for remote access to microdata, aka MONA. With this system users are given secure access to databases at Statistics Sweden from practically any place that can provide Internet access. Data are processed and analyzed through a rich set of applications e.g SAS, SPSS, STATA, GAUSS, Microsoft Office or Super Cross and result sets are then automatically sent to the user's predefined mailbox.

The main goals for the MONA-system are:

- to increase accessibility to microdata for external users at the same time as security and secrecy is reinforced
- to keep all types of microdata for research on site at Statistics Sweden enforcing control of where, when, who and how data are used
- to have instantly upgraded data when needed without any requirements to produce new sets of disks or tapes for redistribution
- to present an easy to use front end for the end users built on well-known standard techniques and components such as server-based computing
- to present a complete system with powerful servers and a rich set of applications with no requirements on expensive equipment and software costs for end users

The MONA-system is built around communication between a client and a terminal server usually called server-based computing. The main idea for this concept is when a client is connected to the server, the client's computer or terminal performs no application processing. It processes only keyboard input and screen output and functions like an input/output terminal. All application processing is done in the server. A non-expensive PC or specialized terminal can be used as a client computer, running any Windows 9x/NT/XP operating system as well as Linux or MAC.

## **6. PROVIDING ACCESS TO MICRO DATA TO RESEARCHERS ABROAD**

In the Nordic countries the same regulation concerning data confidentiality, as for release of data outside the NSIs, are in principle also valid when data is delivered to other countries. There are however some restrictions. According to the Data Protection Directive it is in principle forbidden to transfer personal data that is being processed to a third country (a country outside the EU and EEA) unless the third country in question ensures an adequate level of protection. The Data Protection Acts in the Nordic countries contain similar rules about release of data to a third country. In Sweden the Secrecy Act is also of relevance. According to Chapter 1 section 3, the release of confidential data to an authority or an

international organisation outside Sweden is not allowed unless it is communicated in accordance with special provisions in legislation. Also, the information in a corresponding case might be given to a Swedish authority and the authority holding the information deems it evidently compatible with Swedish interest that the information is communicated. The EU regulation is such special provisions that make it possible to release micro data to Eurostat. There are no other special provisions concerning statistical micro data.

In Sweden the release of micro data to an authority in other countries for research is therefore possible only if it is compatible with Swedish interest that information is communicated. Micro data may be released to private researchers in other countries if it is evident that the information can be disclosed without the person whom the information concerns suffering loss or being otherwise harmed. In practice Statistics Sweden is restrictive with release of de-identified micro data to researchers in other countries. Regards to the Statistical Act in Norway, all users of microdata are bound to secrecy. Since the legislation is not valid outside Norway, and Statistics Norway is thus not able to control if researcher in other countries maintains the confidentiality rules, Statistics Norway find it indefensible to release micro data outside Norway. However, the legislation accept transfer abroad if Norway is subject to an obligation to make a transfer pursuant to an international agreement or as a result of membership of an international organization.

In Finland the same regulations concerning data confidentiality as in Sweden, as for release of data outside Statistics Finland, are valid. An applicant must provide a description about how the data confidentiality is secured in the recipient country. Denmark does not release micro data to researchers in other countries but foreign researcher can use the Danish on-site arrangement under the same conditions as Danish researchers. Iceland has no experience in delivering micro data to researchers abroad.

## **6.1 Eurostat**

The release of information to Eurostat is regulated in the EU regulations on statistics. According to Regulation 1588/903 the national authorities shall be authorized to transmit confidential statistical data to Eurostat. National rules on 3 Council Regulation (Euratom, EEC) No 1588/90 of 11 June 1990 on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities statistical confidentiality may not be invoked to prevent the transmission of confidential statistical data to Eurostat where an act of Community law governing a Community statistic provides for the transmission of such data. This means that NSIs in principle are bound in regulations to release micro data for community statistics. However, transmission of data which are not covered by a specific Community legislative act is voluntary and that national rules can prevent the transmission of confidential data. Transmission of confidential statistical data shall be carried out to Eurostat in such a way that statistical units cannot be directly identified. This does not preclude the admissibility of more far-reaching transmission rules in accordance with the legislation of the Member States.

## **7. FUTURE CHALLENGES**

It is clear that the Nordic countries are committed to improving access to high quality microdata. One important strand in future development is to compile several new thematic registers tailored to better meet the needs of the research community. To accomplish this, considerable work is needed, engaging both methodologists and subject matter experts. Another future trend is to develop techniques that allow linkage of data from different sources, both within and outside the Nordic NSIs. In addition, some of the Nordic NSIs are designing Statistical Data Warehouses, which will enable them to build integrated registers and cubes in ways that allow continuous updates of data. It goes without saying that all these development trends are subject to fulfilment of legitimate confidentiality requirements.

A close cooperation on use of registers and microdata has been launched with some Nordic universities. A number of seminars and symposiums have been arranged and several postgraduate students and a

couple of joint professors are involved in this cooperation, which is expected to grow considerably in volume and importance.

Improved access to microdata involves relatively high costs to be borne by researchers. Because of this, Statistics Sweden has approached The Swedish Research Council arguing that funding from the Council of a system of microdata access would give researchers a lower initial cost when accruing data. This would also facilitate an increasing use of microdata in research. Our view is that the system of microdata access should be regarded as a national facility. Experiences from other areas where basic financing have been arranged and researchers only pay for marginal costs have been very positive. Such a solution would firstly incorporate full IT support for on-line access via the Internet. Secondly, a front office would be installed to serve and advise the researchers involved. Thirdly, this financial support would allow more and better thematic databases, which could be accessed directly by researchers, and could be created at an early stage. A solution along these lines would clearly facilitate an improved access to user friendly, high quality microdata.

## **8. Concluding remarks**

There seem to be good possibilities to improve the access to microdata for researchers and other legitimate users in a radical way without violating confidentiality. In such a statistical system largely based on registers as in the Nordic countries, this really is a major improvement, also bearing in mind the new possibilities for dynamic analysis thanks to longitudinal microdata. Although considerable progress has already been made, systematic work for further improvements pave the way for new opportunities for researchers.

## **References**

MONA Microdata ON-line Access at Statistics Sweden (only in Swedish)