

Distr.
GENERAL

CES/AC.71/2003/23 (Summary)
4 December 2002

Original: ENGLISH

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

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

**ORGANISATION FOR ECONOMIC
COOPERATION AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

Joint ECE/Eurostat/OECD meeting on the management of statistical information systems
(Geneva, 17-19 February 2003)

Topic II: Impact of technical measures and standards on data quality

**MANAGEMENT AND MONITORING OF THE STATISTICAL PROCESS
AND IMPACT ON DATA QUALITY**

Supporting paper

Submitted by the Statistical Office of the Republic of Slovenia¹

Summary

I. BACKGROUND AND NOTES

1. In this paper, we want to describe the possibilities that modern technology offers to national statistical institutes (NSIs) in terms of providing conditions for efficient management and control of processes that have direct impact on the quality of statistical data and services. The impact of technical measures and solutions on the quality of processes is constantly increasing. At the same time, their complexity is growing. We are witnessing substantial changes in the field of IT and NSIs have to be prepared to master the situation and to respond to the new questions and challenges. When we are talking only about technical measures and changes the results are usually quickly obvious, but frequently changes are such that the contents or purpose determine the type and extent of technical measures and solutions, i.e. that new IT initiates new methods of work and enables more efficient work and general improvement of quality. In this context, we will describe the principal measures and solutions in the Statistical Office of the Republic of Slovenia (SORS), especially from the point of view of the impact on the entire production process.

II. GENERAL POLICY AND STRATEGY

2. Quality of data and services must be one of the key strategic objectives. The more complex the processes and the technological infrastructure the greater the need for consistent policy and strategy that determines key obligations and responsibilities of NSIs. A high level of synergy is assumed. SORS's strategy is to reinforce the already established concept of 'Input-Throughput-Output'. Our experience shows that the concept of vertical competence over methodological issues and horizontal competence over infrastructural tasks brings good results. We explain why such a concept is better than the so-called 'stove pipe' concept.

¹ Prepared by Milan Katif (milan.katic@gov.si), Mirko Klanjčič and Pavle Kozjek.

However, to successfully manage and solve possible conflicts within this concept and process the entire life-cycle of statistical surveys and similar tasks – including those that for different reasons do not take place in a standardized way – needs to be defined. An appropriate work plan has to be prepared that contains all deadlines and the names of responsible persons. Also, regular monitoring and control over the quality needs to be established.

III. METADATA CONCEPT AND RULES

3. Metadata are one of the key factors for successful integration of contents on the one hand and processes on the other hand. It is difficult to speak about the quality of official statistics without taking into account the necessary metadata. It is difficult to estimate data quality without the criteria necessary for such estimates. It seems that this is a closely related issue that requires consideration in its entirety. SORS developed a working prototype of a meta database, i.e. a METIS repository, with numerous modules such as: national program of statistical surveys, work plan, publications, questionnaires and questions, statistical variables and characteristics, lists of micro and macro tables in the repository, a module for evaluating the quality, a module for dissemination, etc. A classification server named Klasje was set up, which at the moment contains about 500 standard classifications and nomenclatures and together with the above-mentioned meta database represents the central metadata infrastructure for monitoring real data through the entire production process.

IV. INPUT PROCESS

4. In this part of the paper, we describe the key issues as regards the contents, organization and technology of input processes and placing them in the total production chain or the process of work in SORS also from the point of view of their impact on the final quality of statistical data and services. This reflects in the organizational integration of phases of work in the input function with a special emphasis on transparency and traceability. Recent developments aim at implementing the integral treatment of statistical material (QUESTOR questionnaires) including electronic data collection from primary sources (QonLine) and data capture from administrative sources. The core of the integration of processes and procedures is an integrated input database (SAJPI – Common Administrative Core of Integrated Identifiers). The expected result is the decrease of burden for reporting units, better quality of results and lower costs due to automation of standard procedures. We will deal with the situation and solutions (state-of-the-art) in this field in SORS as well as with standard solutions of classical data input (Blaise Generator) and links with the Metis meta database.

V. THROUGHPUT PROCESS

5. In this part of the paper we describe the most important processes linked with the operation of the statistical data warehouse (DW), which is the general orientation of SORS. Integration of statistical data is made via the unified DW concept, the integral part of which is the Metis meta database. Implementing measures that are linked to data quality dimension is in this segment of crucial importance. Modern concepts emphasize the role of statisticians and IT experts in integration of data with metadata and in setting up a multipurpose database at micro and macro level. In this concept, statisticians and methodologists are responsible for (meta)data in terms of contents, while IT experts do that in terms of technology. The data coming into the throughput segment are administratively clean statistical data and only after statistical control become reference statistical data (FOR - Final Observation Register) and as a rule cannot be put back into the input process. In data access, linking and analysis, the key role is that of Metis and Klasje and common dimensions of the DW. The most important issues in terms of contents, technology and organisation of throughput processes within SORS are treated.

VI. OUTPUT AND PUBLISHING PROCESS

6. In this part of the paper, we describe the key questions and processes linked to dissemination via the Internet. Special attention is focused on quality. These processes have been modernized recently. We describe our concept of technological support that is done on an SQL platform and the PC-Axis family of

tools (i.e. Light edition). Dissemination of statistical tables and data as well as preparation of publications (Web publishing) will be done with the previously mentioned family of tools. However, in addition to this concept, dissemination of certain time series will continue via the Statistical Databank (<http://www.gov.si/bsp/>).

VII. CONCLUSION

7. In the conclusion, we show the most important findings and measures that should be taken into account in individual phases of work with the purpose of providing the highest possible level of the quality of processes, which result in the quality of data and services.
