

**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 UNECE/Eurostat/OECD Work Session on Statistical Metadata (METIS)
(Geneva, 3 – 5 April 2006)

Topic (i) Metadata in a Corporate Context

THE BNSI EXPERIENCE IN METADATA COLLECTION AND ORGANIZATION
Submitted by NSI - Bulgaria¹

I. INTRODUCTION

The National Statistical Institute of the Republic of Bulgaria has begun developing an Integrated Statistical Information System (ISIS). The lack of ensured funding for the complete development of ISIS along with the architecture defined in the Strategy for development of Integrated Statistical Information System and the IT Strategy² requires applying a gradual system of development component by component. The paper presents the approach and organization of ISIS metadata collection at the NSI.

**II. ESSENCE AND GOAL OF ISIS AND PRINCIPLES OF DEVELOPMENT
COMPONENT BY COMPONENT**

1. ISIS development aims at setting up mutually connected object and process oriented information systems using common data bases with controlled access, allowing shared use of the NSI source by experts at the NSI and external experts. A key ISIS component is the Metadata base that is also the basis for the ISIS driving. The architecture of ISIS is defined in accordance with the UNECE Standard 51³. The BNSI top management and experts understand the advantages of the complex ISIS development as an unified integrated system, however the lack of funding is the reason for looking for technological innovation by gradual system development component by component.

2. The gradual development of ISIS could be realized in different ways depending on the consecutiveness in the development of functional and infrastructural components. The advantage of this approach towards a higher technological SIS is conformity with the

¹ Prepared by Svetlana Ganeva, e-mail sganeva@nsi.bg.

² The Strategy for ISIS development and the IT Strategy were developed within the framework of the National Phare Programme project BG 00.06.04 with the technical assistance of Swedish and German experts, as well as the local experts. They are a part of the complex Strategy for Statistics development, which is currently being elaborated and discussed at the BNSI.

³ [Information Systems Architecture for National and International Statistical Offices](#); Statistical Standards and Studies No. 51, Geneva, 1999.

available sources. The disadvantage is the risk of the NSI having to develop a separate project for integration of the already developed ISIS components. The reasons could be the necessity to apply contemporary technology or the specific features of tender procedures under the Procurement law.

3. ISIS development in the BNSI is realized by adhering to the following requirements as a guiding principles:

- The strict adherence to the ideology and architecture of ISIS defined in the Strategy for ISIS development in the work on developing separate ISIS components.
- Common metadata model implementation in the application work on the development of separate ISIS components.
- Providing interfaces between information systems if possible, especially for the metadata exchange. This requirement is compulsory for infrastructural components providing metadata for other information systems-ISIS components.
- The components exploitation must follow developed and approved internal standards.
- The ISIS components must include built in the compulsory data and metadata exchange standards of Eurostat and other international statistical organizations.
- Consolidation of the IT environment and IT elements to be achieved by developing the ISIS components in accordance with the IT Strategy requirements.
- Common development tools to be used in the development of the system.

4. An important requirement for the implemented approach is defining the order for the development of components, while retaining the continuity of the existing statistical information system exploitation without disruption so that every following component be integrated with the already existing ones. This means that the stove-pipe organization of the SIS is maintained. The priority development of the infrastructural components is a possibility to use them in the functional components development. Thus the implemented approach determines the integration of following components to the system until its development in accordance with the approved principles.

5. The BNSI is presently working on development of the following information systems as ISIS components:

- Register of the statistical units (RSU);
- Information system “Statistical classifications” (ISSC);
- Information system “Planning and design of the statistical survey” (ISPDSS).

6. BNSI has developed a bilingual electronic vocabulary of the statistical terms and their definitions to be used internally in the NSI Central office.

III. STANDARDS FOR THE METADATA DESCRIPTION IN INFORMATION SYSTEMS

7. The standard object approach in IS development is implemented through the definition and usage of common objects in all IS. The model used is the MetaNet Reference Model, recommended by EU experts in National Phare project. The BNSI experts discussed the information objects and their attributes one by one and the accepted objects were included in the elaborated technical specifications for information systems (the revision of the technical specifications was required by the contract).

8. The information objects of the ISPDSS are as follows:

- Statistical survey;
- Statistical observation;
- Observation unit (object);
- Global variable;
- Object variable;
- Value domain;
- Question;
- Table;
- Questionnaire block;
- Questionnaire;
- Statistical output;
- Statistical infrastructure task;
- Document;
- Procedure;
- File from an internal source;
- File from an administrative source;
- Population.

These information objects provide the statistical activities description from the planning and design stage to the statistical products production, including design and output of the documents going along with the statistical survey – questionnaire, instructions, methodology, technical specification for the data processing and so on. The information objects attributes allow keeping the link between objects and storing information for the object's history and changes.

9. It was necessary to define new information objects in accordance with the statistical practice traditions in Bulgaria as follows:

- Statistical observation –the national practice shows that it is possible to have one survey realized by more then one observation. At the same time one observation could be useful for more then one survey. By using this information object is possible to describe in a proper way the existing traditions in the statistical survey organization and to manage them.
- Statistical infrastructure task – the information object is intended to collect the meta information for the activities of the statistical office servicing conducted statistical surveys.
- File from an internal source and File from an administrative source –these information objects allow studying the information flows, providing information for the statistical surveys. They are also a tool for the macroeconomic surveys description, which – as a rule – use the information from the statistical surveys (micro databases) for macroeconomic data production.
- There is also an information object „Procedure” intended for procedures description – technological sequence of activities – with a view of the future driving the processes in ISIS.

10. Standard outputs are defined in the ISPDSS, some of them are extremely important for the management of the statistical information system. Such outputs are the National Programme for the Statistical Surveys, the Calendar presenting the results of the statistical surveys and the List of the standard statistical indicators. The ISPDSS provides introduction of data concerning the time and funding needed as a prerequisite for better processes management.

11. The standard for the classifications description in the IS “Statistical classifications” is a combination between MetaNet Reference model and CLASET model. The information objects in the ISSC allow for both the development and maintenance of “standard“ classifications and the establishing and maintenance of the statistical classifications for separate surveys, as well as for extraction of information for classifications used in a definite point in time. It is foreseen to load main classifications used after 1990 in the system. The classifications, the relations between them and the dynamics of their development are described using the information objects as follows:

- Classification family;
- Classification sub-family;
- Classification;
- Classification version;
- Classification variant;
- Classification level;
- Classification item;
- Best classification practice (Case law);
- Table;
- Relation;
- Dynamics;
- Document.

12. Information objects in the Register of the statistical units have an important place in ISIS. Objects in this system include the administrative and statistical units, forming the structure of the national economy. These units are defined in accordance with national and European legislation as follows:

- Administrative units:
 - Legal Unit;
 - Local Legal Unit.
- Statistical units:
 - Enterprise;
 - Local Unit;
 - Kind of Activity Unit;
 - Local Kind of Activity Unit;
 - Enterprise Group.

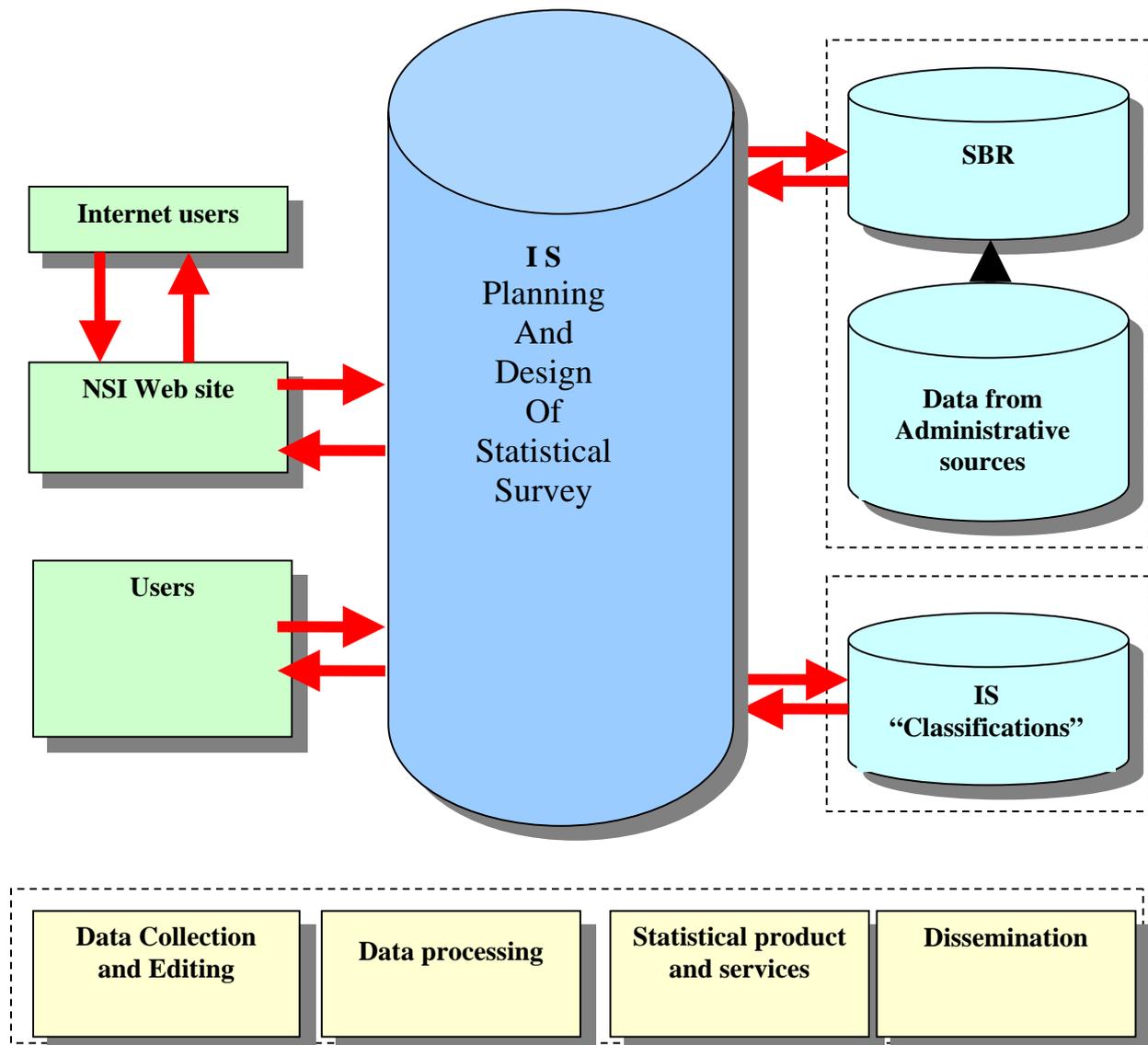
13. The attributes of the statistical units and the statistical units themselves are described in the ISPDSS as the population and sample are defined based on the objects of the Register of the statistical units.

IV. THE PLACE OF THE DEVELOPING COMPONENTS IN THE ISIS ARCHITECTURE

14. The place of the presently developed information system in the architecture of the Integrated Statistical Information System is defined in accordance with their functionality:

- Register of the statistical units is an infrastructural component including the information for the statistical units, their characteristics, their demography and development their “participation” in statistical surveys;
- IS “Statistical classifications” is intended to collect and store meta-information;
- The Information system “Planning and design of the statistical survey” is a functional component developing as an application tool for collection and storage of data describing statistical surveys.

The relations between developing systems are presented by the following scheme:



V. ORGANIZATION OF METADATA COLLECTION – REALIZATION THROUGH THE SI SYSTEMS

15. During the implementation phase of the information system “Planning and design of the statistical survey” it is envisaged to create an organization and through the description of statistical surveys to load information on statistical surveys. Thus the ISPDS becomes an application tool for metadata collection for surveys.

16. Work on the development of information systems at NSI is organized by the project leader, who is exclusively entrusted with task. Coordination is ensured by the Coordination Council, headed by the NSI vice-president and including all directors. This allows work to be monitored by the NSI management and experts on all levels. A task force for the development of each information system was set up, consisting of representatives of all directorates participating actively in the discussion of the relevant issues as well as in the design and development phase and the implementation phase. Project management is done by means of a project management timetable - ISO 9001 for project management.

17. During the implementation phase of the information systems it is envisaged to enter the metadata on statistical surveys carried out by NSI in accordance with a programme adopted by the NSI President. The programme shall be based on decisions taken concerning the consecutiveness of implementation of the data collection and editing component as well as the possibilities for migration of existing microdata databases in survey registers.

VI. FUTURE TASKS

18. Work is in progress and a tender procedure is under way for the development of a Technical specification for development of ISIS as well as on a detailed Technical specification for the development of the Information system “Statistical observation” as a follow-up functional component of ISIS. Based on this detailed specification this functional component has been commissioned as a separate phase design under the contract for the PHARE 2004 National Programme. As a minimum task, it is envisaged to develop the functional component “Data collection and editing” along with the creation of the Observation registers and the Final observation registers as a Microdatabase. The national programme also envisages development of the Data from administrative sources database with a view of creating a database of existing administrative registers usable for statistical purposes with a view of a description of their constituent units and their characteristics and their future integration as component of ISIS. As a next phase, it is envisaged to create follow-up functional components – the processing of statistical data with a view of development of the Macro database and the Statistical products dissemination component, along with development of the Dissemination database as a copy of the Macro database, concerning data which can be disseminated. The consecutiveness of setting up the components of ISIS shall be defined precisely in the technical specification for their development. There are options for developing the Metadata base as a key component of ISIS intended as the integrated system management component. One of these options involves development of the Metadata base and gradually adding to it metadata from newly created components. Another option is to develop the components separately and finally integrate them, followed by the creation of the Metadata base as a common management component. In both cases, the process of building up the components must go in a direction of applying common standards and rules, in view of the possibility for their integration in ISIS.