**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 Seminar on the Management of Statistical Information Systems (MSIS)
Sofia, Bulgaria, 21-23 June 2006

Topic (i): Changes in statistical processes

INTEGRATION OF INFORMATION SYSTEMS IN COMPONENT BY COMPONENT DEVELOPMENT OF ISIS – THE BNSI PRACTICE

Invited Paper prepared by Svetlana Ganeva et al., National Statistical Institute of Bulgaria[1]

I.   INTRODUCTION

1.   The National Statistical Institute of the Republic of Bulgaria has begun to develop an Integrated Statistical Information System (ISIS). The lack of ensured funding for the complete development of the ISIS along with the architecture defined in the Strategy for Development of the ISIS and the IT Strategy[2] requires applying a gradual approach to the development of the system component by component. This paper presents the actions aimed at integrating the components of the future integrated system, which are being developed as separate information systems, into the ISIS.

II.   ESSENCE AND PURPOSE OF THE ISIS AND PRINCIPLES OF ITS DEVELOPMENT COMPONENT BY COMPONENT

2.   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

---

[1] Prepared by Svetlana Ganeva, e-mail sganeva@nsi.bg and Pavlina Paskaleva, e-mail ppaskaleva@nsi.bg. In the report the authors have used the internal standard Requirements for Information Systems Integration, prepared by the company-developer Acsior-Lirex in cooperation with the NSI experts.

[2] Within the framework of the 2000 PHARE National Programme, project BG 06.04.00, a Strategy for Development of the ISIS and an IT Strategy have been elaborated with the assistance of Swedish and German experts, as well as local experts. They are a part of Strategy for Statistics Development, which is currently being elaborated and discussed at the BNSI.

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 51[3]. 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.

3.   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 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.

4.   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.

5.   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.

6.   The BNSI is presently working on development of the following information systems as ISIS components:
   ➢   Register of the Statistical Units";

---

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

> Information System "Statistical Classifications";
> Information System "Planning and Design of the Statistical Survey";
> Information System "Demography";

7. The NSI also has an Electronic Glossary of Statistical Terms for use within the framework of the NSI.

III. TECHNOLOGICAL REQUIREMENTS TO THE INFORMATION SYSTEMS IN THE PROCESS OF THEIR DEVELOPMENT

8. The ISIS systems at the NSI are being developed as connected components that can be integrated and through which a prototype of the future ISIS is being implemented. During the development of the systems, we have determined and carried out the solutions that will ensure not only the implementation of each of the systems separately, but also their integration. These solutions include:
> standardised object-oriented approach;
> common functionality of the systems;
> internal system services.

The application o these solutions to the other components will ensure the gradual development of the ISIS.

9. The BNSI practice in the development of information systems can be summarised as follows:
> Requirements for integration of the information systems are determined and approved as a standard for development of the ISIS components. By means of these requirements, standardised information and technological resources for common usage are determined, which will be provided by the implementation of the separate subprojects (information systems);
> During the development of the components "Planning and Design of the Statistical Survey", "Statistical Classifications" and "Register of the Statistical Units", a unified metadata model is applied and when necessary and possible, the information objects are common;
> IS "Statistical Classifications" provides a service for the other components – the service provides a tool for development of classification variants for the statistical survey and extraction of the classifications necessary for functioning of the information systems;
> The components "Planning and Design of the Statistical Survey" is a software application for description of a statistical survey applying a standard for description of a statistical survey – there are opportunities for description and storage of the metadata about the statistical survey in a structured form with a view of their use in the development of the functional components and metadata base in its integrity as a key component, by which the processes in the ISIS will be driven;
> In the process of description of the statistical survey, the procedures of conduction of the statistical survey are also described in a standard manner, owing to which a unified approach to metadata collection in a structured form for the entire process of collection, processing and conduction of the statistical survey is applied.

IV.   TECHNOLOGICAL ASPECTS OF THE INTEGRATION

10.   The standardised object-oriented approach includes the object presentation of the information and the processings in accordance with the Object Management Group (OMG) principles and standards. The information objects, as well as the operations with them are separated in the following groups:

(a) Information objects for common usage with read-only access:
   – classifications, maintained by the Information System "Statistical Classifications";
   – metadata about the statistical surveys, maintained by the Information System "Planning and Design of the Statistical Survey";
   – legal units, enterprise, local unit, kind-of activity local unit, kind-of activity unit, group of enterprises, maintained by the Information System "Register of Statistical Units".

11.   The information objects for common usage are maintained by (are within the technological scope of) one IS, but they can be used simultaneously and independently of other information systems. A prerequisite for common usage of information objects is their stable and in practice constant structuring and organisation of the data and relations for a long period of time, as well as the opportunity every information system to have its view about the objects, which is derivative or compatible with their presentation in an IS maintaining the respective objects. The organisation of the use of objects by an IS has a general and specific part. The specific part is implemented by each separate IS. The general part is compulsory for all IS. As matters stand, the objects for common usage are created, modified and deleted by one IS and are used "for reading" by other IS. In case it is permissible certain characteristics (and/or relations) of the object to be created and modified by other IS as well (IS–users in the terminology above), a policy on rights and conflicts solving is determined and provided, which is implemented within every IS–user. The diagram below presents the use of common information objects. It is simplified and shows the access to information objects maintained in the information base of one IS (IS 3) on the part of other information systems (IS 1 and IS 2). This access can be direct using the concrete physical presentation of the object in the information base of IS 3. It can also be indirect using relevant standard techniques for creating a view about the information object different from its immediate physical presentation. These standard techniques include:
   - defining and using View objects in the information base of IS 3;
   - defining and using objects through which "standard packing" of the physical object in the database is done: EJB;
   - using XML descriptions for "representation" upon physical objects from the databases and standard interpreters for their use, etc.

IS 1

IS 3
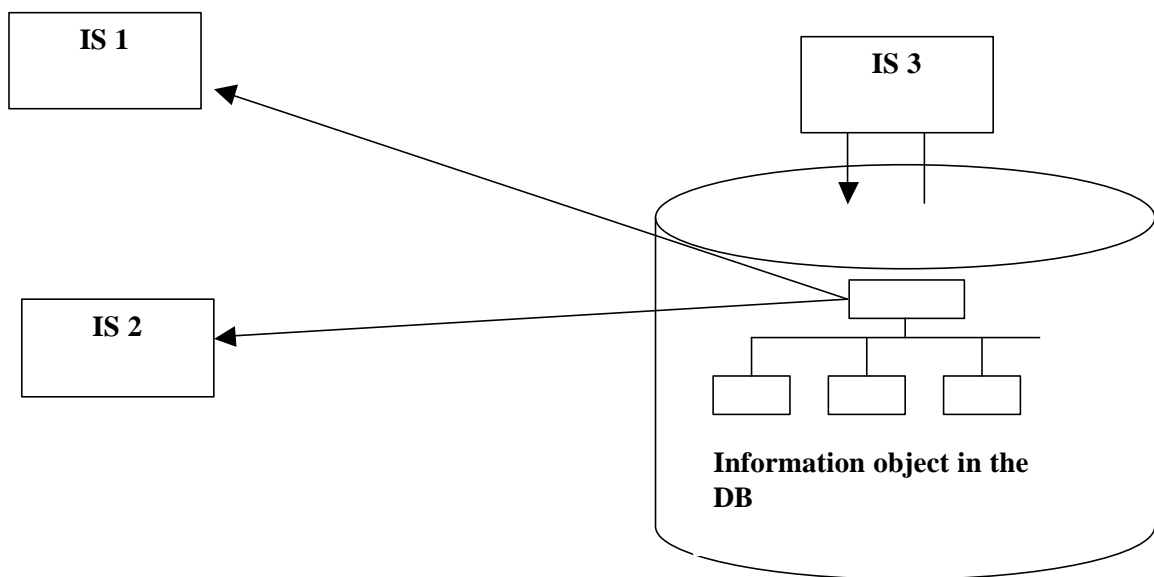
IS 2

Information object in the DB

Diagram1: Use of common information objects

12.   The use of one-way arrows indicates that this method of use of the information objects is intended mainly for "read-only". The responsibility for maintenance of the data about the object is born by the IS in whose information base it is located.

(b) **Common information objects:**
- NSI's structural units (organisation chart);
- employees;
- users;
- groups of users (roles);
- resources – system objects for which access rights are granted (screen forms, menu items, etc.);
- granted access rights;
- configuration parameters of the system;
- definitions of messages of the system;
- record of attempts at security violations;
- record of the processings;
- context of the activity performed;
- signal message;
- message for an error.

13.   The functionality of the systems is presented as an aggregate of technological events. The organisation of the use of the systems related to the regimes of work will be accomplished ensuring unification of the processings in all regimes. This approach ensures parameterisation and external adjustability of the systems with regard to their functions. The processings in the systems are parameterised. The context of the processings, the main elements of the formal logical control of the information being entered, signals and control functions of the systems,

correspondence of the functions and work places are administered by externally adjustable structures of data.

14. The common functionalities are identical for more than one IS and because of this they are realised as common "components". These components are included in the project and in the application of each IS which uses them and the respective IS uses the realisation of the common functionality "such as it is". The conditions for integration of the common components in the applications, as well as for realisation of specific functionalities that belong to the subject matter of the common functionality, but are not included in it because they are concrete for certain separate IS, are created in the projects of the respective IS. In case the component with the common functionality requires or performs exchange of messages with other components of the same IS (elements of WorkFlow, etc.), the interaction with it is carried out through interfaces. The interfaces are defined by the component for common usage and are implemented by each IS which uses it. Examples for common functionalities are: "Maintenance of context"; "Control of the access to resources"; "Records of "logical transaction"; "Error processing"; "Service of systems' own All System Information (ASI)"; "Registration of an attempt at security violation"; "Generation of a signal message"; "Sending a message to/through an external service", etc. The diagram below shows the use of the common functionalities:
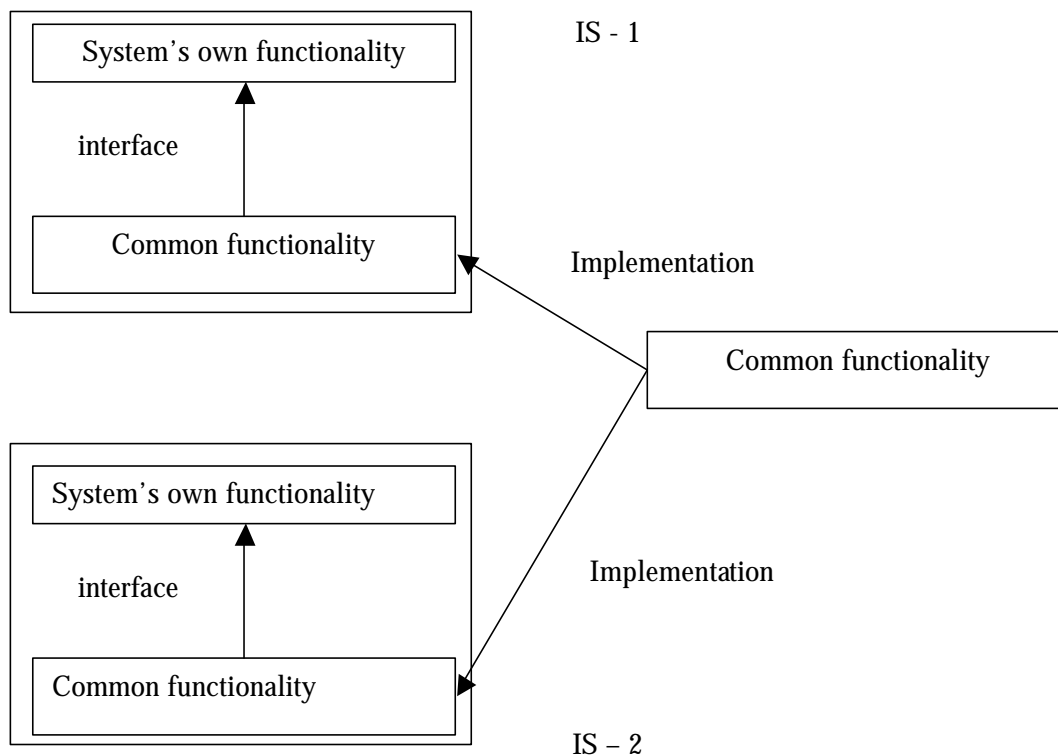

Diagram 2: Use of the common functionalities

15. Internal system services – in the implementation of the application software, a major role plays a large set of internal services ensuring the use of the unified processings, all system information and security control. The services are functionalities that are entirely within the

scope of one of the four IS, but they are used by the rest IS. The organisation of the service is entirely within the subject matter of the respective IS.
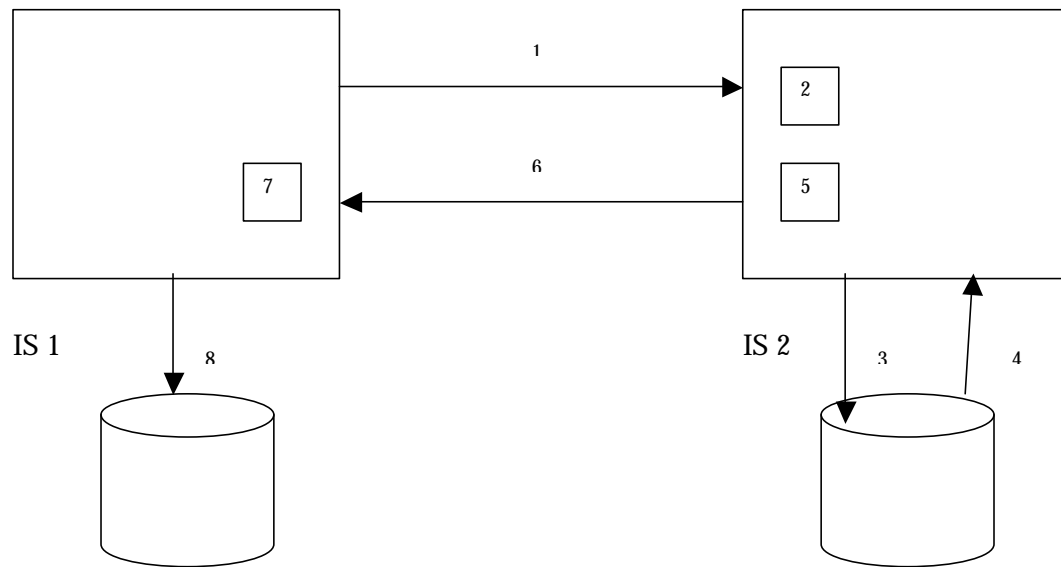

Diagram 3: Use of a service

Symbols:
1.   IS 1 places an order for a service provided by IS 2;
2.   IS 2 takes the order for a service and determines the legality of the order and conditions for its execution;
3. & 4. IS 2 uses data from its own information base to execute the service;
5.   IS 2 forms the result of the service for which an order is placed;
6.   IS 2 provides IS 1 with the result of the executed service;
7.   IS 1 receives the result of the executed service;
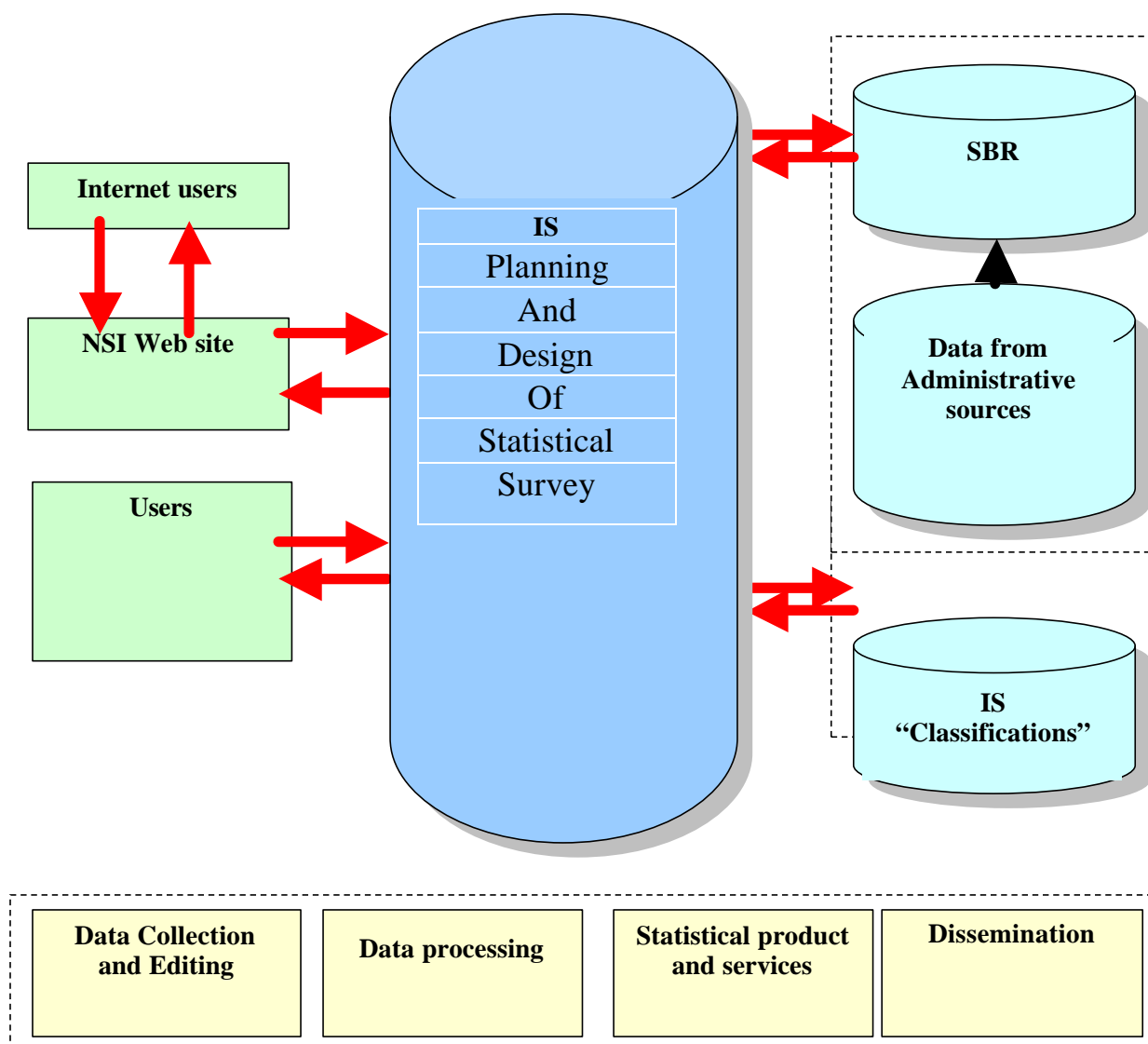8   IS 1 enters the result of the service execution in its own information base.

16.   IS "Statistical Classifications" provides the services that will be accomplished within the scope of the present requirements:
- Provision of a classification:
  - General specification of the "parameters" of the service:
  - identifier of the classification;
  - history;
  - specification of the result: way of provision, type of result, contents;
- Provision of a single item of a classification:
  - General specification of the "parameters" of the service:
  - identifier of the classification;
  - way of identification of the item;
  - identification value;
  - history;
  - specification of the result: way of provision, type of result, contents;
- Provision of subordinate meanings in a hierarchy:
  - General specification of the "parameters" of the service:
  - identifier of the classification;

- way of identification of the main meaning;
- identification value;
- history;
- specification of the result: way of provision, type of result, contents;

- Provision of a main meaning in a hierarchy:
    - General specification of the "parameters" of the service:
    - identifier of the classification;
    - way of identification of the subordinate meaning;
    - identification value;
    - history;
    - specification of the result: way of provision, type of result, contents;

- Provision of changes in a classification:
    - General specification of the "parameters" of the service:
    - identifier of the classification;
    - time period;
    - specification of the result: way of provision, type of result, contents.

V. THE PLACE OF THE COMPONENTS THAT ARE BEING DEVELOPED IN THE ARCHITECTURE OF ISIS

17. The diagram of the components that are being developed currently and the available systems allows their integration in the following way:

| Data Collection and Editing | Data processing | Statistical product and services | Dissemination |

VI. ORGANISATION OF THE WORK ON DEVELOPMENT OF ISIS COMPONENTS

18.   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 ISPDSS becomes an application tool for metadata collection for surveys.

19.   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.

20.	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.

## VII.	FUTURE TASKS

21.	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.