

**STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

**COMMISSION OF THE
EUROPEAN COMMUNITIES**

CONFERENCE OF EUROPEAN STATISTICIANS

EUROSTAT

**Joint UNECE/Eurostat Work Session
on Statistical Metadata**
(6 - 8 March 2002, Luxembourg)

Working Paper No. 23
English only

Topic (iii): Metadata and quality

**ORGANIZATIONAL STRUCTURE OF METADATA IN
THE STATE STATISTICAL COMMITTEE OF AZERBAIJAN**

Submitted by the State Statistical Committee of Azerbaijan¹

Contributed paper

I. INTRODUCTION

1. The role of information and statistical information in society has increased with the transition from a planned economy to a market economy. The number of users of statistical information has hugely increased. Requirements of information sources have also changed. Today, users want to obtain data in a more convenient form, with a detailed description of parameters.

2. In previous years, the only users of statistical information were the state bodies; now, our information interests the various commercial organizations, businessmen, private individuals, educational institutions and also international organizations.

3. The information site of the State Statistical Committee (SSC) of Azerbaijan on the Internet was posted at the end of 1997. The information on the site is submitted as static tables, their updating is burdensome and is thus not undertaken on a regular basis. Such an approach does not meet modern requirements. Therefore, the question of the integration of modern information technologies in SSC is of great importance.

II. BASIC CONCEPTS

4. At the present stage of information technologies development, the collection, processing and distribution of information in many respects depends on the correct performance of the information, i.e. on correctly prepared metadata.

5. In the report on "Development of information technologies in the State Statistical Committee of Azerbaijan", all aspects of information technologies development are considered, including the creation of an information database.

6. In this report prepared by experts on information technologies of the SSC, information bases are considered from the aspect of macro-, micro- and meta-databases as input and output databases.

7. Bearing in mind the multilevel structure of the SSC, where output data for one level is entry for the following level, microdata is simultaneously input and output data. The structure of microdata is often subject to changes (the structure varies and there is a variety of reports). In some cases the microdata also may be represented for research purposes for various establishments.

¹ Prepared by Mehman C. Ibrahim.

8. Two aspects of macrodata are considered: the macrodata for the top level (for leading bodies) and the macrodata for dissemination. It is necessary to bear in mind the role and quality of metadata in collecting, processing and distribution of statistical information. With the help of the qualifying standards IMF, UNSC and ECE used for developing and creating metadata, and the experience of the international statistical agencies in SSC, we have developed projects for micro, macro and meta databases.

9. In the report on "Development of information technologies of the State Statistical Committee of Azerbaijan", the various qualities (which fully comply with IMF and ECE standards) to be applied to metadata at SSC are briefly considered:

- ?? to help the user search the required information (should contain the data definition, a level of aggregation, the key data for search, the logical menu of search, interfaces between various parts of an information system, etc.);
- ?? to help with definition of the required information (understanding of value, periodicity and limitations of the data, correct interpretation and value of the data, etc.);
- ?? to help with a rating of the required information (reliability and quality of the data in detail, methodological aspects connected with the data in all stages of processing, etc.).

III. DEVELOPMENT OF THE INFORMATION SYSTEM

10. In line with the project to develop information technologies in SSC, the following must be created:

- ?? meta database of metadata;
- ?? meta database for micro databases of different levels of aggregation (fourth and third levels – these are statistical offices of administrative regions of the Republic, district statistical departments of Baku and regional statistical offices of Nakhchivan AR, second level - the statistical office of Baku and Statistical Committee of Nachichivan AR and first level – the departments of central office and corresponding departments of HCC);
- ?? meta database for macro databases (the same meta information could be used for the interpretation of macro data for a higher level - government bodies - and macro data for dissemination).

11. While working on the project of a database for metadata, we considered two variants for the creation of an informational structure of metabase, the second of which is preferable for us:

- ?? creation of metadata for every database on every level: fraught with the creation of special catalogues, this scheme complicates the search and finding of required bases;
- ?? creation of meta database for metadata for micro and macro databases on every level: this should facilitate the search and finding of required information bases and required information and should nullify the necessity to create a special reference manual accompanied.

Based on the above, those meta database projects already finalised must now be modified somewhat.

12. The last two years have seen work on the creation of databases. The methodology to be used to create and administrate the metadata and database is not finalized, and is, in fact, at the revision stage. This methodology will be completed after some experience is gained in the peculiarity of our system and the technology and software used.

13. The metadata project for Azerbaijan was planned according to the standards of IMF, UNSC and ECE. The materials from the Joint UNECE/Eurostat Work Session on statistical metadata, and also pilot metadata from Kazakhstan were used. The Kazakhstan metadata was obtained with the assistance of IMF specialists sent to SIC countries for this specific purpose.

14. A special group consisting of heads of SSC departments and chiefs of functional departments of HCC led by the deputy director of HCC has been created to assist in accelerating the work and in obtaining the necessary materials from departments of SSC.

15. During this period, some work was undertaken collecting the necessary information for classifying and projecting meta database for macro database (for high level government bodies and macro data for dissemination) and the project for pilot metadata of Kazakhstan was launched. The first version of the project of metadata base (for high level dissemination) may be modified for full compliance to the qualifying standards of international organizations (especially the dissemination aspect).

16. Operations are underway to collect the necessary information for classification and designing of a meta database for microbases and aggregated databases at various levels in the structure of SSC.

IV. METADATA BASE OF SSC OF AZERBAIJAN ON INFORMATION DISSEMINATION

17. A metadata base consists of meta information on databases by branches of statistics for the top level and information distribution. Below there is shown an exemplary structure of a metadata base.

Name of metadata fields	Specification of metadata fields
The title of the object	The name under which the given object is known
Identifier	Simple, the original reference to the object. The reduced name of a database, it is possible to use an abbreviation. Should be semantic loading. Necessarily by letters of English alphabet.
Description of database	It is shown at viewing, and should be in English and national languages. A choice of language will be made by the request of the user.
Complete description of database	It is shown at viewing, and should be in English and national languages. A choice of language will be made by the request of the user.
Server	The name of the server, where the database is located.
Path	Path to the database.
Database	Name of the SQL database.
Type	Describes the type of the variable: ? -metadata base
Administrator	Name and occupation of administrator.
Date	Date of creation or last modification of the base.
Version	Number of version.
Language	The descriptions of parameters should be submitted in several languages (Azerbaijan, English, Russian).

18. The base will include the description of the register of enterprises, the following classifications - administrative – territorial units, kind of activities, types of ownership, goods and services, units of measurement and also branch data bases - national accounts, financial and bank statistics, industry statistics, consumption, agriculture, trade, demography, welfare, labour statistics, price statistics, transport and communication statistics, education, environment, population census, etc.. The main network administrator is responsible for the design and change of the structure of the main base.

V. METADATA FOR BRANCH DATABASES AND TABLES OF GOSKOMSTAT OF AZERBAIJAN ON INFORMATION DISSEMINATION

19. The branch base consists of metadata on a database or tables on a given branch of statistics. Below is an exemplary structure of a metadata base.

Name of metadata fields	Specification of metadata fields
Name of Database	The shorten name of a database, it is possible to use an abbreviation. It should have a meaning burden. Compulsorily by letters of English alphabet.
Description of database	It is shown at viewing, and should be in English and national languages. A choice of language will be made by the request of the user.
Complete description of database	It is shown at viewing, and should be in English and national languages. A choice of language will be made by the request of the user.
Category	Describes the access type to base: public or special.
Server	The name of the server, where the database is located
Path	Path to the database
Type	Describes the type of the given variable and might take following values: ? - metadata, ? -Database, ? -Table, R-register, V-variable.
Database	Name of the SQL database.
Timescale	Period of updating the information in base.
SpecialProcBefore	If it is necessary the name of the special procedure is given.
SpecialProcAfter	If it is necessary the name of the special procedure is given.
Administrator	Name and occupation of administrator
Date	Date of creation or last modification.
Version	Number of version.

20. Exemplary structure of metadata of tables.

Name of metadata fields	Specification of metadata fields
Name of main base	Shortly name of main (connected) data base, might be used abbreviation. Should be semantic loading. Necessarily by letters of English alphabet.
Name of table	Shortly name of table, coincide with the name of table in connected main base.
Description of table	It is shown at viewing, and should be in English and national languages. A choice of language will be made by the request of the user.
Complete description of table	It is shown at viewing, and should be in English and national languages. A choice of language will be made by the request of the user.
Category	Describes the access type to base: public or special.
Server	The name of the server, where the database is located
Path	Path to the database
Type	Describes the type of the given variable and might take following values: V-variable.
Database	Name of the SQL database.
Timescale	Period of updating the information in base.
Variable	Name of column in a table.
Valueset	Name of dataset in table
SpecialProcBefore	If it is necessary the name of the special procedure is given.
SpecialProcAfter	If it is necessary the name of the special procedure is given.
Administrator	Name and occupation of administrator
Date	Date of creation or last modification.
Version	Number of version.

21. Design and change of the structure of the branch base and tables is carried out by the data base administrator upon agreement with the main network administrator.

VI. THE CONCLUSION

22. The structures given here are preliminary. In SSC of Azerbaijan, active work in the study of best practices in information technologies and opportunities separating RDBMS and OS are conducted. **As the decision on creation of corporate network SSC is accepted on the basis of OS Linux there are studied opportunities RDBMS MySQL and also an opportunity of use RDBMS Oracle under Linux.** The main reason for the backlog in SSC in the creation and introduction of databases and application of modern technologies of information distribution is the lack of opportunity to purchase the appropriate technologies (software and hardware). The real structure and meta information will be submitted in the appropriate databases.