

**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 (ii): Metadata Concepts, Standards, Models and Registries

**RECENT ACHIEVEMENTS IN METADATA SYSTEMS AT THE ROMANIAN NATIONAL
INSTITUTE OF STATISTICS**

Supporting Paper

Submitted by National Institute of Statistics, Romania¹

I. INTRODUCTION

1. In the framework of Romanian National Institute of Statistics there are some domains in full swing. Among there we can discern the follows: Intra-Community Trade INTRASTAT, Central Database TEMPO and Special Data Dissemination Standard SDDS.

2. These new domains require a special metadata approach. In terms of each topic, the used metadata have different parts. The specific of mention domains claims a real challenge for metadata management.

II. THE INTRASTAT METADATA

3. According to Romanian accessing to European Union, National Institute of Statistic is constrained to assume two different ways for building his foreign trade data collection. The first one will be the registration of intra - Community transactions, an activity that will be carry on NIS. The second one will be the uploading of data about extra - Community trade, according to the records provided from Romanian Customs Authority.

4. In the first case is necessary that NIS will design a register which contents data about Romanian exporters and importers. This register must be considered one of the main components of INTRASTAT metadata system. A government decision established that exporters - importers register will be starting from data furnished by Romanian Customs Authority, in the first year of this statistical survey, and Ministry of Public Finance, in the future, according to information content in the new format of VAT questionnaire.

¹ Prepared by Edmond - Lucian SINIGAGLIA, lucians@insse.ro.

5. Starting with data entrance, the metadata will play a significant role. The transaction will be registered if the input data will be filtered according with metadata related. Therefore will be verified the existence in nomenclatures and registries of the following information: commodity code (according to Combined Nomenclature and List of goods excluded from statistics relating to the trading of goods between Member States to be transmitted to the Commission), partner country (EU countries list), flow (export/import), details about transaction nature (according to custom regime), mode of transport, coding of delivery terms (Incoterm ICE/ECE Geneva).

6. Another metadata used in the data uploading is the interval system. Depending on the types of goods, the range of value is defined in relation to the weight and/or in relation to supplementary units. That gives us the following possibilities (an interval for the value/weight, an interval for the value/supplementary unit, an interval for the weight in kg/supplementary units). There are any cases in which the intervals are fixed by the Combined Nomenclature itself.

7. In contrast with extra - Community trade data, which will be load exhaustively from Customs Authority, the INTRASTAT data will be process in the NIS and the role of metadata will be increase.

III. THE CENTRAL DATABASE TEMPO

8. The Central Database with Time Series TEMPO permits the management in an integrated conception the NIS main data thesaurus and the metadata related. Through the integration process with metadata, the statistical information can be dominated in a good way. Also the data misinterprets are minimized.

9. In TEMPO system the data are stored in multi-dimensional matrixes or hyper-cubs. If a matrix have many dimensions, her values are most detailed. Each matrix dimension has associated a statistical nomenclature or classification.

10. Every data matrix has two compulsory dimensions: the time reference and the unit of measure in which are expressed the matrix values).

11. The data matrixes are classified into a hierarchical structure which has many levels:

- sector;
- domain;
- sub-domain;
- indicator.

12. Each matrix has associated certain metadata categories, such as:

- the indicator definition;
- information about used nomenclatures/classifications;
- methodological items (range, adjustment methodology, break into time series);
- information about data sources;
- information about data confidentiality.

13. The main terms used in TEMPO, joined with data and metadata categories, are described in glossary of items, which is accessible through main menu of application.

14. The database TEMPO consultation is possible owing to a client type program, designed in Delphi C/S 4.0. The access to data and metadata related is made on account of Oracle database, being of NIS server.

15. The first option offered in the main menu is SEARCH, which contains the following options:
 - CATALOGUE;
 - NOMENCLATURE;
 - MATRIX CODE;
 - KEY WORDS;
 - PERIODICITY;
 - DATA SOURCE;
 - ADJUSTMENT;
 - LAST SEARCH;
 - LOAD FROM DISK;
 - MATRIXES LIST IN MEMORY.

16. It is obvious that the data access is facilitated by a set of metadata. It is possible to select a data matrix starting with the elements of a nomenclature/classification, such as:
 - a nomenclature subset;
 - a position contents in a subset/entire nomenclature.

17. There are any facilities for selecting/deselecting of the nomenclature/subset nomenclature.

18. There is the possibility to retrieve the data starting to key words, which can be looked as a sort of metadata. These metadata, which can be classified as retrieval metadata, are:
 - the indicator name;
 - the indicator definition;
 - the matrix name.

19. VISUALIZATION, another option of main menu, provides other facilities, among other things a lot of metadata:
 - EXTRACTING SPECIFICATION;
 - DATA;
 - METADATA;
 - GRAPHICS.

20. Choosing the METADATA option, the system shows the metadata categories associated to current matrix, divided into many areas, such as:
 - DEFINITION;
 - TIME INFORMATION;
 - METHODOLOGY;
 - DATA SOURCES.

21. In the framework of metadata visualization, there is information related to:
 - periodicity:
 - period of start;
 - the last period loaded;
 - period of end.
 - absence of values;
 - record keeping data;
 - successive matrix;
 - number of loaded values:

- definitive;
- temporary;
- setting right.
- break into time series:
 - the break reasons;
 - the break period;
 - conversion algorithm;
 - matrix with compatible data.
- adjustment:
 - methodology of adjustment;
 - associated adjusted matrix.
- other methodological notes.

IV. Special Data Dissemination Standard SDDS²

22. In accordance with the national Strategy for Development of Statistics Romania subscribed to the Special Data Dissemination Standard SDDS in May 2005. SDDS is established by the International Monetary Fund to guide member states that have or might seek to have access to the international capital markets in the dissemination of economic and financial data.

23. SDDS, in taking a comprehensive view of the dissemination of these data identifies four dimensions of data dissemination:

- the data (coverage, periodicity, and timeliness);
- the access by the public;
- the integrity of the disseminated data;
- the quality of the disseminated data.

24. The IMF's Dissemination Standards Bulletin Board (DSBB) presents the components of the SDDS concerning:

- the metadata;
- the methodological skills;
- the quality and accessibility, equality of data-users and thus emphasizes transparency in the compilation and dissemination of the official statistics.

25. The responsibility for the accuracy of the metadata, including timely updates, and for the economic and financial data underlying the metadata rests with the subscriber.

26. The National Summary Data Page (NSDP) includes a wide range of macroeconomic indicators (data categories) related the follow sectors of the national economy:

- Real Sector;
- Fiscal Sector;
- Financial Sector;
- External Sector;
- Population.

² A part of this section is presented in a form received from IMF.

27. In relation to the NSDP the SDDS requires the elaboration and maintenance of an Advance Release Calendar (ARC) that comprises provisional information concerning the timing of the data releases for a four months horizon.

The metadata section of SDDS contains:

- Information by data category (select information by accessing a list of data categories);
- Information by country (select information by accessing a list of subscribing countries);
- View information about metadata dimension and metadata elements by data categories and countries;
- View information by key concepts within for one or more metadata elements countries and data categories;
- Advance Release calendar (ARC) information for one or more data categories and countries;
- View summary of observance information: cross - country practices versus the SDDS.