United Nations ECE/CES/2011/50

Economic and Social Council

Distr.: General
1 June 2011

English only

Economic Commission for Europe

Conference of European Statisticians

Fifty-ninth plenary session
Geneva, 14-16 June 2011
Item 3 of the provisional agenda
Organization of data collection and sharing, and the management challenges for the implementation of Statistical Data and Metadata eXchange

The management challenges of Statistical Data and Metadata eXchange implementation: Israeli Central Bureau of Statistics as a case study

Note by the Israeli Central Bureau of Statistics

Summary

The Israeli Central Bureau of Statistics (ICBS) is basing more and more of its statistics generation on administrative data sources. Those sources serve as an upgraded alternative process for the direct data collection in the field as well as a potential source of information for producing more and of better quality statistics. SDMX has been chosen to act as a leading concept in which the idea of comprehensive data infrastructure will be promoted. This paper describes the managing attitude to the project, the approach to the widthwise data infrastructures method, and the technical tools that were already developed for the use of the end users.

I. Work concept - comprehensive data infrastructure

1. The Israeli Central Bureau of Statistics (ICBS) is basing more and more of its statistics generation on administrative data sources. Those sources serve as an upgraded alternative process for the direct data collection in the field as well as a potential source of information for producing more and of better quality statistics, to be disseminated locally and internationally. Since those processes are among the main business processes of the organization, the right data infrastructures are required.

2. The goals for creating those data infrastructures are:

(a) Implementing a unique methodology within ICBS units;

(b) Preventing waste of resources caused by repeating the same process in different places;

(c) Adjustment of definitions to international standards and regulations;

(d) Adopting the concept of generic work procedures within the ICBS;

(e) Uniformity in publications.

3. The Widthwise data infrastructure is based on four elements:

(a) Dictionaries and code lists;

(b) Metadata standards;

(c) Statistical Data and Metadata eXchange (SDMX) standard;

(d) Data Bank.

4. While each of the elements is significant, the combination of them provides an extra benefit, when considering the following:

(a) The SDMX concept contains references to each of the other three elements as they are all mandatory for the implementation of SMDX;

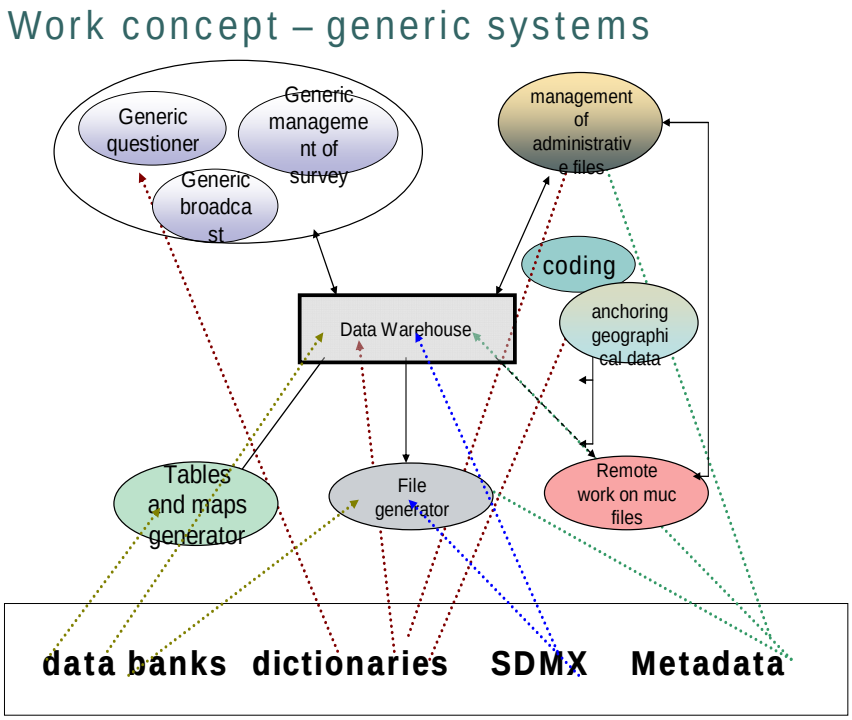
(b) These elements are addressed as part of the work concept established by the institutes leading the subject.

5. The SDMX has been chosen to act as a leading concept in which the idea of comprehensive data infrastructure will be promoted.

6. According to this concept, the management of the SMDX in the ICBS is not just exchanging data in a certain standard but building all the relevant systems that support and are the base for the SMDX operations.

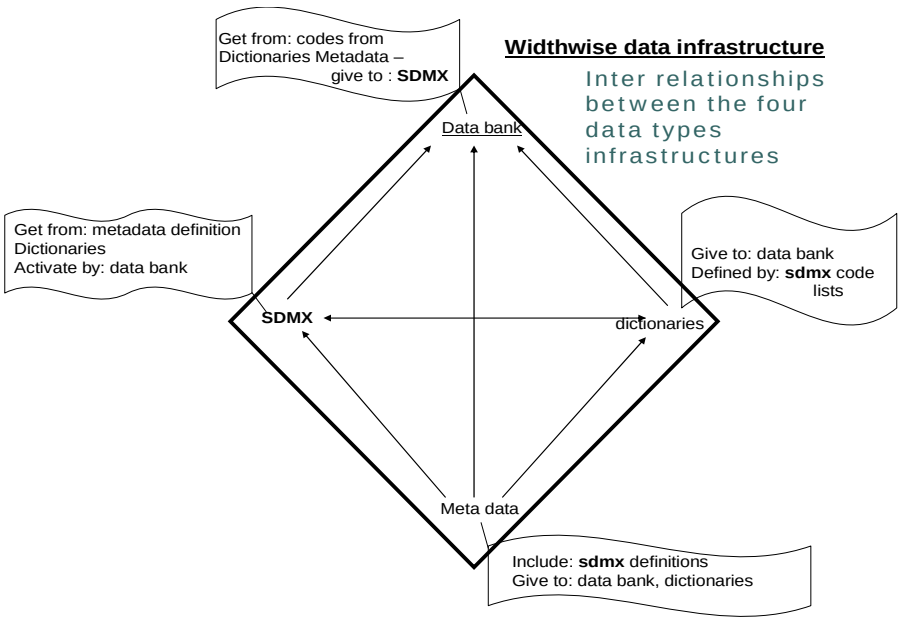
7. The concept of the SMDX and the comprehensive data infrastructure also integrate the technological work concept of the ICBS. A concept supports developing generic systems for the benefit of the bureau's activities which are diverse, as comprehensive infrastructure is generic by definition.

8. The following is a diagram that demonstrates the concept of generic systems, the place of the comprehensive data infrastructure and the SDMX in that concept.



II. Methodology, definition and management

9. The methodology of promoting the SDMX concept is derived from identifying the relationship between different infrastructure's systems. Accurate identification, analysis and relationships definition will simplify by definition the processes of gathering the data and disseminating it, which becomes mostly technical. In the following scheme the relationship between the widthwise data infrastructure elements and the SDMX are presented:



10. For an accurate definition of the relationship between those infrastructures there is a need to define the business process that creates the data which is the source for the data exchange, according to the SDMX standard.

11. The purpose of this definition is managing the organizational internal statistical process while processing the data and creating the time series according to the SDMX standards.

12. Managing the methodology means:

(a) Specifying the flow chart between the infrastructure systems;

(b) Specifying the flow chart between the infrastructure systems and the processing and dissemination systems.

13. In addition, each system is managed separately, by defining rules for: assembling and updating the code lists, the metadata content and its management, definitions and management of the database and for the definition and management of the data retrieval for the SDMX, according to the required format and standard. A very central element is working according to international regulations and integrating them into the internal work process of the ICBS.

III. Technology, supporting systems and their management

14. In order to promote the methodological concept there is a need for technological tools which will form the base of the framework of the work process. These tools can be divided into groups:

(a) System for assimilation and managing the code lists / dictionaries. A basic system was developed for the 2008 population census and currently it is being adapted to the SDMX regulations. In the system's concept there is a content expert for each subject (Education, Demography etc.) and a system administrator who is in charge of the development. In the near future the system is going to receive the code lists that were

defined for the pilot that the ICBS is involved in with the OECD, using export of services data;

(b) Defining Data Scheme and Data Flow Model according to SDMX standards for the pilot regarding the export of services data. The scheme was built on the basis of an existing scheme and with the required adjustments to the SDMX standards. Stating the obvious, in the future the statistical database scheme should be built from scratch, according to the defined regulations;

(c) System for the Metadata management - assimilation, archive and distribution are to be built;

(d) System for the data dissemination in SDMX formats, including definitions of data and metadata structure definitions, and the required xml's. A prototype was built for the pilot.

15. In any future work concept the ICBS should build web services which allow the local and international databases to synchronize directly and automatically. This vision – machine to machine connections - will be possible only when all the preliminary systems will be ready.

IV. Implementation

16. The management of the implementation of the work concept and the use of the technological tools demands demonstration of the added value of working in SDMX environment. This added value is on two levels:

(a) Organizational level. Unified work processes, unified code lists which will enable multi discipliner cooperation and upgrading the quality of data, correspondence of the reports according international regulations, building of generic process which will save time and money, and more;

(b) The level of each section. In this level the added value derived from saving work time in reporting the international organizations by using the machine to machine method, uses of existing templates and systems for the benefit of each business process, which implies reduction of development and maintenance, the ability to manage statistical procedures, to supervise them, and more.

V. Summary

17. For achieving the goals of the SDMX approach, the following procedures should be implemented:

(a) Establishment of comprehensive data infrastructure and supporting tools for their management;

(b) Reorganization of the database on the basis of the updated scheme;

(c) Feeding the database from the comprehensive data infrastructure and pulling the data from it, according to the SDMX definitions and with the help of the supporting tools.

18. Until today the ICBS worked by this concept at the pilot which took place with the Organisation for Economic Co-operation and Development (OECD). However, the supporting groups were not generic, but temporary tools.

19. In a holistic point of view for the all bureau the SDMX method should be assimilated by the right order of processes and systematically for getting a generic work process that will serve either the ICBS while producing the data and those organizations who expect to get this data from the ICBS.