

Distr.  
GENERAL

CES/SEM.47/5 (Summary)  
9 January 2002

Original: ENGLISH

**STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE**

**COMMISSION OF THE  
EUROPEAN COMMUNITIES**

**CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROSTAT**

**Joint UNECE/Eurostat Seminar on Integrated Statistical  
Information Systems and Related Matters (ISIS 2002)**  
(17-19 April 2002, Geneva, Switzerland)

Topic I: Application of web technology to integrate statistics

**METADATA REQUIREMENTS FOR THE INTEGRATION OF A DECENTRALIZED  
STATISTICAL SYSTEM AT THE OECD**

**Invited paper**

Submitted by the Organization for Economic Cooperation and Development<sup>1</sup>

**Summary**

1. The OECD is an intergovernmental organization, with 30 member countries, which has as its objective to help member governments achieve their goals. The structure of the OECD is very much similar to that of governments in member countries and is divided by subject matters such as agriculture, environment, economics, social affairs, etc. Each of them has a statistical unit which collects and compiles statistics for its ongoing work. Although there is also a central Statistics Directorate, mainly responsible for the bulk of economic statistics and for internal and external co-ordination in the area of statistics, the internal organization of statistics at the OECD is for a major part decentralized.

2. The decentralized structure of statistics makes statisticians and analysts work closely on each subject, which creates a virtuous circle of quality improvements. However, working across subjects is made difficult because of the lack of common standards and tools. Presenting statistical information in a corporate way in a decentralized structure is a challenge that current IT technologies can help addressing. However, there are a number of prerequisites in terms of cooperation tools and metadata to help the integration of statistical activities. The OECD is putting together the information necessary to seamlessly integrate the various components of its decentralized statistical system and to help collaboration between individual statistical units.

3. The following objectives have been identified:

- ❏ Increase the collective knowledge about our statistical information; make that knowledge accessible;
- ❏ Develop a glossary of statistical terms;
- ❏ Increase data standardization and metadata management; and
- ❏ Develop corporate data access via Reference Series and a common "browser".

---

<sup>1</sup> Prepared by Gérard Salou (gerard.salou@oecd.org).

4. A project has been launched to collect information on individual statistical activities and to share it amongst statisticians and other users of statistics. This includes high-level information on the nature and objective of each statistical activity, the people involved in, technical information on systems used for collection, data management, and outputs. Plans for the future are also collected. This is the highest level of the corporate information system. This collection as well as its maintenance requires wide participation of statisticians in the Organization.

5. A common classification of statistical activities has been agreed upon throughout the Organization which has now been implemented. Consequently, datasets can also be classified using the same classification. This will be used to help users navigating the OECD collection of datasets.

6. In order to facilitate the location of information through definitions, a glossary of statistical terms has been developed. It contains a comprehensive set of target definitions for the main variables collected by the Organization for use in its statistical and analytical activities and publications. It draws extensively from existing international statistical guidelines and recommendations. The definitions contained in the OECD Glossary are, in particular for those relating to variables collected by the OECD, “target” definitions based on existing international statistical recommendations and guidelines. National definitions used in the actual compilation of data by OECD member countries may (and frequently do) depart from these standards for a number of reasons. Information on national definitions, concepts, etc., for specific data collected from member countries is normally collected by subject-matter areas.

7. The glossary provides a highly visible and readily accessible definitional information for use by OECD author areas in the development of questionnaires and other data collection instruments, for reference in databases and for inclusion in published output. The Glossary is intended to facilitate the collection of consistent data by the various Directorates of the Organization. The Glossary is also intended to lend greater transparency to OECD data requirements from national agencies that provide information to the Organization. This set of target definitions will ultimately be linked to actual data located on OECD databases.

8. OECD databases contain a massive quantity of statistics. Most of them are to be used in the area of expertise where they are maintained. In general, individual statistical areas only need a few elements of data coming from other statistical areas. Those data have been identified and are designated as “Reference Series”. A virtual database and a common browser are in development to permit an efficient access to Reference Series with no need of specific expertise or technical equipment.

9. Finally, the development of a common data catalogue has been started. In a first phase it will contain references to the most important data series used in the Organization, In the long term all data used throughout the Organization will be referenced there.

10. In summary, this cooperation effort aims to provide users with seamless and efficient access to OECD statistics from different sources in a decentralized organization. It is a complementary effort to the IT development of a common interface (the “Common Browser”). Other benefits are the identification of areas of data duplication and improvements in the quality of OECD statistical data and metadata. Finally, the development strategy described in this paper leads to the development of a coherent statistical information management architecture for the OECD. The paper concludes by presenting OECD’s experience as a model for the integration of statistics in e-government initiatives.