

Distr.
GENERAL

CES/SEM.47/25 (Summary)
16 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 IV: Ways of making statistical information systems more responsive to users

**IMPROVING ACCESS TO STATISTICAL INFORMATION AT OECD
IN RESPONSE TO USERS' REQUIREMENTS**

Invited paper

Submitted by the Organization for Economic Cooperation and Development¹

Summary

1. The OECD is recognized throughout the world for its reliable and comprehensive analytical and statistical work. The gathering and harmonization of international data in a multidisciplinary environment are key to international comparison and policy making. The rapid and easy access to statistical information is a major challenge for many organizations, particularly if the management of certain data collections is decentralized in specialized subject matter areas. This paper focuses on work carried out by the OECD Secretariat to render its statistical data and metadata accessible for different constituencies.
2. Presenting statistical information in a corporate way while being managed in a decentralized structure is a challenge that modern ICT² can help to address. Traditionally, decisions on the organization of statistical processes have been taken giving priority to the requirement of individual statistical units. This has resulted in efficiency gains within individual statistical activities at the expense of the co-operation between analysts from different subject matter areas for cross subject studies.
3. A vision for an OECD wide statistical data and meta-data environment has been developed based on a survey of user needs. The main focus of this vision is to provide a unified view of OECD statistics through a common data browser, comprising a repository of reference series, common search and query tools as well as a central catalogue with comprehensive information about the various data sources.
4. The survey highlighted the importance of the following needs for internal users of OECD statistics:
 - Locating OECD 'Reference' series (most popular series - i.e. Exchange rates, CPI, GDP, Unemployment);
 - Locating data by theme from different sources;
 - Global keyword search across directorate for specific series or across an entire catalogue;

¹ Prepared by Peter Lübker, Head of Systems Development and Support (peter.lubkert@oecd.org).

² Information and Communication Technology

- Making data more visible/accessible to across directorates;
- Clarifying/highlighting areas of duplication;
- Monitoring data access.

5. The OECD has been working on innovative ways to offer new services for the management and dissemination of statistical data and metadata. First steps were taken in 1996 with the development of SQL based “Corporate Data Environment” (CDE) to offer Intra- and Internet access to dynamically updated data sets through a catalogue from standard Web browsers.

6. More recently, efforts have concentrated on improving the on-line access to a broader range of official OECD statistics independent of their source database environment. The interactive access to these sometimes rather large data sets has complemented the more traditional electronic publishing channels/media. An important consideration in this context was to ensure that the respective data sets are compiled only once for use on different media. This not only allows a reduction of the amount of preparatory work and time to market, but also diminishes the risk of inconsistency. As with the CDE, all data is stored in a single, common format.

7. In line with the OECD Chief Statistician’s³ “New Strategy for the OECD Statistical System”, a variety of technical solutions for implementing a “Common Browser” to facilitate the access to statistical information internally have been studied. In light of the already available experience with the CDE architecture the main focus was on technology that would permit to build a “broker function” between different source (or production) databases. Another important technical consideration was that relational database management system vendors are now integrating OLAP⁴ functionality into their database engines. After some prototyping and with the help of a complementary study of state-of-the-art tools in this domain, a data-warehouse approach is now being considered. This concept, which involves ETT/ETL⁵ tools for the regular updating of the warehouse, will be applied to a set of key reference series used across all Directorates of the Organization. Main design and implementation considerations are as follows:

- Ability to search across and locate data by theme (a central data catalogue and dictionary);
- Complete metadata (descriptions, sources, usage, last update, etc.);
- Ability to save selected data in a format suitable for export to analysis software of the user’s choice (Excel, E-Views, FAME, etc.);
- Dynamic updates of production databases (to avoid data duplication);
- Adoption of a scalable approach: other data sets could be added later;
- Leverage existing developments (CDE, STD Browsers, OecdFAME Wizard, etc.).

8. Robustness, security, and maintainability are key advantages of a data warehouse approach over the concept of a database “broker” which retrieves data from different sources on the fly. A data warehouse allows to completely separate changes in the production environment(s). It also allows for simpler information update and control procedures.

9. In summary, this development aims to provide users with seamless and rapid access to key OECD statistics from different sources using a common interface (the “Common Browser”). Other benefits cover savings in staff time (training, data management), the identification of areas of data duplication, and improving the quality of OECD statistical data, especially metadata. Finally, the development strategy described above is intended to serve as a building block in the development of statistical information management and services architecture for the OECD. The objective is to federate “vertical” sets of data and metadata into comprehensive repository through a crosscutting catalogue, a glossary of statistical terms, and an overarching database of OECD statistical activities.

³ Enrico Giovannini - OECD Chief Statistician and Director of the Statistics Directorate.

⁴ Online Analytical Processing.

⁵ Extraction Transformation Transport/Extraction Transformation Language.