

**Distr.
GENERAL**

**CES/AC.71/2001/RP.1
30 January 2001**

ENGLISH ONLY

**STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

**COMMISSION OF THE EUROPEAN
COMMUNITIES (EUROSTAT)**

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE/Eurostat Meeting on the Management of Statistical Information Technology
(Geneva, Switzerland, 14-16 February 2001)

Topic (i): The impact of data warehousing on the management of statistical offices

TOPICS FOR DISCUSSION

Submitted by Lars Rauch, Statistics Sweden

ROOM PAPER

1. Will the implementation of data warehouses require a reengineering of the statistics production process? What has to be changed, if necessary?
2. What are the consequences for the management caused by the movement from a stovepipe organised production to an integrated production approach for the statistical production system;

The development of output databases and data warehouses may change the paradigm of the statistics production process that traditionally is subject area oriented. We can also observe changes in the pattern of data collection using new technologies as Internet and new methods to reduce the burdens for respondents. Should not such important changes have an impact on the kernel statistics production process? Would it be worth to investigate whether the statistics production process (aggregation, tabulation, analysis) really should remain stovepipe organised?

3. Data consistency – the main problem in a data warehouse implementation? How can sophisticated metadata systems support the data consistency?

The integration of statistics in a data warehouse approach leads to demands for a much higher degree of data consistency that has to be ensured by metadata, in particular by a harmonised and standardised classification system. Statistical metadata are more and more used in the production process but still there is a lack of unification and office-wide harmonisation. A central managed metadata system, as a centralised metadata resource could be the basis for a significant improvement of the situation. But that is again mainly a management problem to put together the different interests in a large organisation as NSO's usually are.

4. Management of disclosure risks in a data warehouse environment

The dramatically extended possibilities to combine data from different sources, drill down functionalities, cross tabulation features, etc. will have a strong impact on disclosure issues. That cannot be left only to subject experts and IT staff. In the first line it is a task of the management to create the necessary awareness of the problems and to manage the preconditions for efficient solutions.

5. Will it in future be necessary also to reengineer the delivery of data to international organisations?

The development of data warehouses not only in NSO's but also in international organisations will probably also influence the way to deliver data – in both directions. It maybe a question in the future to investigate whether the present way to deliver data still is the most efficient. Why not to deliver cubes in the future? What about the metadata? Are more metadata necessary as we stress it for the dissemination and presentation of statistics on the national level? Can XML in connection with data warehouse technologies be an alternative?

6. How to organise the development, implementation, administration, and maintenance of a statistical data warehouse?

In contrary to the stovepipe organised statistics process, the development and maintenance of data warehouses are certainly not a responsibility of a particular subject matter department. That is also already true for centralised databases. But the implementation of data warehouses will require a higher degree of cooperation. It is obviously a management task to organise such a development. Harmonisation and unification of processes that are distributed in the organisation will not work without a strong and efficient management.

7. The impact of the Internet on the dissemination policy using a data warehouse;

8. The impact of the data warehouse technology on the web sites;

The Internet is going to be (in some cases it already is) the main channel for presentation and dissemination of statistics for a NSO and also for international organisations. Using data warehouses as a basis for the access by users will require new functionalities and methods. On the one side the data warehouse creates demands on functionality to be provided via Internet. But also the fast development of the Internet technology as such will offer new possibilities that can be used by statistical Internet websites. From the management side it is necessary to observe that development and to draw conclusions for the own work.

9. What are the main advantages the users of statistics will gain from the usage of data warehouse technologies as a dissemination vehicle?

The user of statistics should always be in focus of the activities of a NSO. The development of data warehouses must lead to significant advantages for the users. The data warehouse technology applied in an efficient way will provide a faster and more flexible way to access data. It should also make the data access much easier. The presentation methods should include flexible graphical and geographical visualisation. That will lead to a new quality of statistics presentation.