

STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE  
CONFERENCE OF EUROPEAN STATISTICIANS

Working Paper No. 2/Add.6  
English only

Work Session on Statistical Metadata  
(Geneva, Switzerland, 18-20 February 1998)

Item 6 of the provisional agenda

**COMMENTS ON  
“STANDARDS FOR STATISTICAL METADATA ON INTERNET”**

Submitted by OECD

GE.98-

1. A group of experts from the OECD Statistics Directorate met to discuss this document. Their comments and suggestions are noted below. This group was composed of: Rita Varley, Head of Main Economic Indicators Division; Gérard Salou, Head of the Statistical Technology Section; Pierre Beziz, Administrator in Main Economic Indicators Division.

**General comments:**

2. We understand that these "Standards for Statistical Metadata on the Internet" are proposed for National Statistical Offices (does it also include national agencies disseminating official data?), not for International Organisations. Our position in commenting on this document will then be that of user of statistics and not that of provider of statistics.

3. Page 2, paragraph (i) *Internet is a common means . . .* We agree that different users need different levels of detail for the metadata. However, we think that additional items should be included in the "minimum set of metadata", such as the Definition of the indicator for instance.

4. For the indicators published in our monthly "Main Economic Indicators", we publish two types of methodological information, the level of detail in each being different:

(i) Summary information: "Sources and Definitions" provides brief up-to-date explanations on the sources and definitions used by the Member countries and the OECD Secretariat, to compile indicators. The metadata items described are: Source, Definition, Coverage, Collection and Calculation.

(ii) Very detailed information : "Sources and Methods" provides more detailed methodological information describing how indicators are compiled. See "List of Metadata Items for OECD's Main Economic Indicators" (paper presented at the UN/ECE Work Session on Statistical Metadata, 22-25 October 1996, Berlin) for the detailed list of metadata items.

5. The five metadata items described in our Sources and Definitions can serve as a starting point for an agreement on a "minimum set of metadata items" for the "general public".

6. Page 3, paragraph (iii) "*Properly organised metadata . . .*": We think that keeping the metadata up-to-date is fundamental. We suggest that all Internet pages incorporate a time-stamp (date of the last update).

7. Page 3, paragraph (iv) "*Homepages of statistical . . .*": We also think that common guidelines for organising the homepages may help some countries to "fill the gaps" in their own methodological description, and reach standards for the compilation and dissemination of statistical information.

8. Page 3, page 4 paragraph (v) "*Metadata and structured organisation . . .*": We fully agree that access to data and metadata should be organised hierarchically. We also think that data and metadata should be integrated, that is, metadata should be accessible directly from data.
9. Page 5 paragraph "*Metadata supporting clients . . .*": It is not clear for us what "data at the variable level" means. Calendar of release dates and names of contact persons constitute very useful information and should be made available on the Internet.
10. Page 6 paragraph "*Metadata for interpreting . . .*": The Statistics Directorate of the OECD does not collect microdata, so we have no special experience in this area. We think that an agreement on a standard list of metadata items for macrodata should be reached first: this could be the objective of the next meeting.
11. The items you propose (for information on macrodata) should be expanded and developed. We think it is very important for the correct interpretation of the data to describe other items such as the Data Collection procedures (reporting units, reporting methods, survey details, etc.). In the case of the Unemployment series, for instance, the "general users" should know for a correct interpretation whether the data are derived from administrative records (register unemployment) or from Household surveys.
12. Our list of metadata items (points 1 to 5; see Annex) could serve as a starting point.
13. Page 7 paragraph "*Metadata to assess the quality . . .*": Point number 2 is not directly linked to data quality. It should be described in the data collection procedures (see previous paragraph). Methodological information concerning points 1, 3 and 4 is essential to assess data quality. Information for points 5, 6 and 7 is very detailed and may be grouped under one single item "Other elements on data quality".
14. Page 8 paragraph "*Some general considerations . . .*": We find that all the points mentioned here are extremely relevant. As you suggested, the user's Internet access time should be saved, for example by: reducing the number of pictures and images, making a table of contents available, and implementing tools to facilitate site navigation.
15. We hope that these comments and suggestions will help in proposing minimum standards for metadata on Internet.