

CONFERENCE OF EUROPEAN STATISTICIANS

UN/ECE Work Session on Statistical Metadata
(Washington, D.C., United States, 28-30 November 2000)

Topic (iii): Needs and responsibilities of international organizations for metadata

STATISTICAL METADATA AND THE IMF: LEVERAGING AN ONGOING INVESTMENT

Submitted by the International Monetary Fund¹

Invited paper

1. The launch of the International Monetary Fund's Data Standards Initiatives, beginning with the Special Data Dissemination Standard (SDDS) in the Spring of 1996 and continuing through the establishment of the General Data Dissemination System (GDDS) in December 1997, marked the beginning of a period of intense metadata development for many statistics producing agencies including national statistical offices, central banks, and Ministries of Finance. Although statistical metadata similar to that called for under the SDDS existed for some countries in various formats prior to the SDDS, the standardized format and comprehensive coverage of SDDS metadata presented challenges for many subscribers. As a consequence, subscribers to the SDDS and, more recently, participants in the GDDS, have invested substantial time and effort in developing their metadata for posting on the Dissemination Standards Bulletin Board (<http://dsbb.imf.org>). Moreover, maintaining these metadata in a current state also requires significant resources. The challenge, therefore, is to leverage the benefits arising from these investments for the benefit for a wide range of data and metadata users without imposing added burdens on metadata providers.

I. ONGOING INVESTMENT IN METADATA

2. The format for the presentation of SDDS metadata on the DSBB uses four dimensions or perspectives to describe a statistical series; these are: a) a description of the coverage, periodicity, and timeliness of the data; b) information on access to the data including advance release calendars; c) information on the integrity of the data including the terms and conditions under which the data are compiled and disseminated; and d) information that sheds light on quality such as the identification of methodological publications and related data series that serve as a cross check on the data. For more information on the requirements of the SDDS, see <http://dsbb.imf.org/overview.htm>. In addition, SDDS subscribers provide detailed descriptions of data compilation practices called summary methodology statements; most of these statements are prepared according to a standard outline, or set of "prompt points," to encourage uniformity across subscribers.

3. Countries that participate in the GDDS seek to develop further their statistical systems and provide metadata similar to that of the SDDS, augmented by plans for improvement. Given that these metadata also are posted on the DSBB, the GDDS provides a framework to assist participants, international and regional organizations, and aid donors in better coordinating their technical assistance efforts. As of September 12, 2000, 11 countries have their metadata posted on the DSBB, and this number is expected to grow steadily as countries see measurable benefits accrue to GDDS participation.

¹ Prepared by Robert Di Calogero.

4. Subscribers to the SDDS are required to update their metadata as needed on a timely basis. This means not only providing timely changes to so-called “housekeeping” metadata such as the names and electronic addresses of contact persons and changes in the titles of publications, etc., but also changes to the descriptions of compilation and dissemination practices. The aim is to provide one-to-one correspondence between the data series described on the DSBB and the data disseminated on so-called National Summary Data Page (NSDP), i.e., web sites maintained by the subscriber containing the actual data prescribed by the Standard. While metadata updates should be provided continuously as required, the Standard also provides for quarterly certification of the accuracy of the metadata by subscribers, thus ensuring that at least four times a year subscribers scrutinize their metadata to ensure that it is timely, accurate, and complete.

5. While the provisions for maintaining timely and up-to-date metadata under the GDDS are not as specific, participants are expected—as a minimum—to review and update their metadata on an annual basis. While more frequent metadata updates for GDDS participants are encouraged, limited resources for statistics in most participating countries make frequent updates difficult.

II. USAGE OF THE DISSEMINATION STANDARDS BULLETIN BOARD (DSBB)

6. Usage of the DSBB has grown steadily since its establishment in September 1996. In its first three years the number of unique users and number of total hits per month grew from 3,800 and 25,000 in 1996 to 22,600 and 148,000 in 1999. However, beginning in the Spring of 2000, growth in activity has accelerated to an average of 19,400 unique users and 640,000 monthly hits. This sharp growth in average monthly hits has been due to a number of factors, in particular the richness of the metadata on the DSBB as more summary methodology statements have been posted and the addition of a number of new features including, inter alia, the Week Ahead Calendar and the Data Quality Reference Site. Also, there has been a significant increase in the number of hyperlinks from the DSBB to subscribers' NSDP, thus making the web site a “one stop” shopping source for both data and metadata.

7. The sharp increase in use of the DSBB can be interpreted as a significant return on the investments of SDDS subscribers in the development and maintenance of high quality metadata, which undergirds their data dissemination practices. It also reflects, to a more limited extent, innovations in the DSBB to provide products for users, such as the Week Ahead Calendar, that draw upon the metadata. Following the results of a study of the DSBB by outside consultants in 1999, the Fund plans to introduce a number of improvements to the infrastructure of the DSBB to make it a state of the art web site. These improvements will include a content management system for metadata based on Extensible Markup Language (XML) and the introduction of enhanced search and query facilities for users based on the XML environment. These improvements should further increase the value of existing metadata to users and provide the basis for the development of the DSBB as a portal web site linked to other sites containing metadata in a common XML based language.

III. METADATA SHARING AND ANCHORING

8. So far, the discussion has emphasized the substantial investments made by subscribers to the SDDS and participants in the GDDS in the development and maintenance of statistical metadata, and how these investments have been yielding significant returns, as measured by increased use of the DSBB. We also have briefly noted how the introduction of new Information Technology can greatly increase the value of these metadata to users in the future. What remains is to discuss briefly the role the DSBB can play in the future in the broad arena of web sites maintained by other metadata providers and compilers, in particular international and regional organizations. In considering this question, it is useful to consider the impact of the Fund's data standards initiatives—especially the SDDS—in terms of the following issues, namely: (a) concerns on the part of organizations that redisseminate metadata, that they not overburden providers of such information with repetitive requests; and (b) the varying formats in which metadata has been provided.

9. First, because the SDDS is a standard, subscribers have committed to maintaining their metadata in a timely and comprehensive fashion. This commitment is reinforced through the maintenance of the NSDP, which should contain exactly those data series described in the metadata. Also, participants in the GDDS have pledged, as a minimum, to annual reviews and updates of their metadata.

10. Second, since its introduction in 1996, the metadata model or format of the SDDS has been adopted, to some degree, by other organizations, thus contributing to the development of a common platform for metadata presentation and analysis.

11. Therefore, given the commitment to timely and comprehensive metadata that is part of the SDDS, and, to a lesser degree, the GDDS, and the increased use of the model or format for metadata presentation, the DSBB can play a useful role as a reference point or anchor for statistical metadata. Discrepancies between metadata described on the DSBB and in other sources for the same data series could be usefully raised with the contact person named on the DSBB and copied to the Fund. Although, given subscribers' commitment to maintain current metadata under the SDDS, users/compiler may assume that the SDDS metadata would always be the most current, differences can arise with other metadata sources for a variety of reasons and the DSBB would seem to be a logical place to initiate an inquiry. The usefulness of the DSBB as an anchor would be constrained, however, by the focus of the Fund's data standards initiatives on a limited number of data categories.

12. Over time, the role played by the DSBB can be greatly enhanced as a portal web site linking the metadata provided to the Fund with metadata compiled by other international and regional organizations. This endeavor, as well as the ability of other metadata websites to link with the DSBB, will depend on the extent to which a common language for statistical metadata in XML can be developed and adopted by various organizations. Work towards the latter could be pursued through a consortium of interested organizations and the Statistics Department of the Fund stands ready to play a role in such an effort.