SDMX: Moving Forward

The heads of statistics of the six organizations that sponsor the Statistical Data and Metadata Exchange (SDMX) initiative met on June 14, 2002 to agree on concrete projects that will advance their joint initiative.[1] The institutions and their respective constituencies favor open and effective international cooperation in maintaining and if necessary developing new e-standards for their data exchanges. New standards should take advantage of the new web-based technologies and the expertise of those working on the business requirements and IT support for the collection, compilation, and dissemination of statistical information.

Four projects, three of which bring "under the SDMX umbrella" work that is already ongoing in several institutions, have been proposed and agreed in outline, and work on them is planned to move forward immediately.

- A practical case study on emerging e-standards for data exchange

- Maintaining and advancing existing standards for time series data exchange

- Creation of a common vocabulary for statistical metadata

- Development of a framework for metadata repositories

Further information on the projects is given below. Participation in them is open to all interested parties. To learn more about them or about SDMX in general, please visit the SDMX web site at www.SDMX.org.

A PRACTICAL CASE STUDY ON EMERGING E-STANDARDS

The objective of this project is to investigate the practical ways in which emerging web technologies can be used to standardize and simplify the collection, compilation and dissemination of statistical information and to identify the core set of e-standards that are required to achieve the full potential of the emerging web technologies (e.g., XML and web services).

The project will serve as a reality-check on possible opportunities by demonstrating just where these new techniques may offer improvements and advances over current ways of doing business, now and in the medium-term.

The project will use the Joint BIS-IMF-OECD-World Bank External Debt Statistics as a practical case study. These data are currently collected from creditor sources through different international statistical reporting systems and are already made available to a global user community over the Internet. They are expected to be complemented in the second half of 2003 by national statistics based on debtor sources.

---

[1] The sponsors are the BIS, ECB, EUROSTAT, IMF, OECD, and UNSD. See www.SDMX.org for the common statement by these organizations.

The case study will be closely coordinated with developments in other SDMX projects and the knowledge gained over the course of the project will be used to formulate proposals for future SDMX work.

The project will be divided into three tracks: identification of key data-exchange models and scoping of demonstrations, experimental standards-setting, and practical demonstrations. Work on all tracks will occur in parallel.

Project deliverables (timeframes)

- Detailed project plan (August 2002)

- Identification of key models and scope of demonstrations (December 2002)

- Experimental standards for exchange and sharing of statistical information (1Q 2003)

- Demonstration of case study examples (2Q 2003)

- Proposals for next steps (3Q 2003)

MAINTAINING AND ADVANCING STANDARDS FOR TIME SERIES DATA EXCHANGE

All SDMX sponsoring institutions have agreed to support and use the GESMES/CB profile of the UN/CEFACT Generic Statistical Message (GESMES) as their de facto standard for batch time series data exchange.

The aim of the project is to maintain GESMES/CB for existing users and to promote GESMES/CB among potential users, particularly where this is to the benefit of SDMX organizations. The project will build on the cooperation that has taken place so far among the SDMX organizations and representatives of users' institutions in the GESMES/CB Maintenance Working Group. The project will develop more formal open structures and procedures for the maintenance and development of the GESMES/CB standard, based on best practices for standard-setting at the international level. The approach developed will serve as a model for the maintenance of other potential e-standards that may be developed under the SDMX umbrella.

The project will promote GESMES/CB through increased publicity and support (e.g., documentation, training, and the sharing of expertise).

Finally, the project will take on board the issue of an XML syntax for GESMES/CB.

Project deliverables (timeframes)

- Detailed project plan (August 2002)

- Develop business case and cost/benefit analysis for implementation of GESMES/CB in XML (October 2002)

- Draft standards-setting process for GESMES/CB (December 2002)

- Draft Quick Start Guide to GESMES/CB, to serve as an introduction to GESMES/CB for first time users (December 2002)

- List of institutions and organizations that are (or are interested in) using GESMES/CB (December 2002)

- Proposal concerning the use of XML syntax for GESMES/CB (December 2002)

- Revised draft of the GESMES/CB implementation guide (October 2002)

- Promotion program, including a strategy for training, to foster adoption of GESMES/CB on a world wide basis (March 2003)

A METADATA COMMON VOCABULARY (MCV)

The objective of this project is to develop a common vocabulary of key metadata items that describe statistical concepts and methodologies used by statisticians in the collection, processing and dissemination of statistical data. Such a common vocabulary (MCV) should be consistent with existing international statistical guidelines and recommendations, and with the terminology being used within SDMX organizations and national statistical agencies.

The envisaged product will take the form of a draft vocabulary of key metadata items, related standard definitions, and context explanations. The vocabulary will be used in existing dissemination frameworks, such as the IMF's Special Data Dissemination Standard and other frameworks used by national statistical agencies. The product would reside in an electronic glossary on the Internet linked to the SDMX web site.

Project deliverables (timeframes)

- Detailed project plan (July 2002)

- Identification of standard metadata items that need to be defined, including those required by other SDMX projects, and the collection of appropriate definitions from existing sources for each of those items (September 2002)

- Review of draft vocabulary by members of the SDMX project, and identification of possible sources of definitions to fill in gaps where definitions derived from international standards are not available (September 2002)

- Public consultation with other organizations and interested parties (December 2002)

- Incorporation of feedback into first production version of MCV. The set of definitions will be available in a glossary, accessible through the SDMX web facilities (Q1 2003)

METADATA REPOSITORIES

The objective of the project is to facilitate and broaden the exchange of metadata among international agencies and across countries through standardization of the organization of repositories (or collections) of metadata.

The immediate objective is twofold: identify the commonalities in the metadata structures of macroeconomic datasets that are collected and stored in repositories; and use these commonalities to develop standardization. Work will take into account the experience with

the development and maintenance of metadata repositories by the IMF (e.g., Dissemination Standards Bulletin Board) and EUROSTAT (e.g., Euro Indicators).

The long-term objective is to have searchable web sites of national and international authorities and private sector organizations using the same search vocabulary. A single inquiry could then be sent to multiple sites that would allow compilation of aggregate results. Here too there will be synergies with the case study as well as with the other projects focused on data and metadata models and vocabularies.

Project deliverables (timeframes)

- Detailed project plan (August 2002)

- Business model for macroeconomic metadata repositories accessible through the Internet that enables access to such information (December 31, 2002)

- XML structure for the rendering of the macroeconomic metadata resident in these repositories that will employ the Metadata Common Vocabulary defined by the MCV project. (December 2002)

- Business model of the acquisition/dissemination of the information within these repositories (June 2003)

- Protocol for the acquisition/dissemination of the information in these repositories implemented in XML (June 2003)

PARTICIPATION IN SDMX PROJECTS

SDMX intends to foster broad participation on a world-wide basis for the projects it is sponsoring. To accomplish this, SDMX will develop working procedures, such as the following, that allow voluntary participation by all interested parties in the public and private sector.

- Build on the work of experts involved in emerging e-standards, including international standards organizations and consortia

- Post draft deliverables and progress reports on the SDMX web site and soliciting public comment

- Directly alert institutions, organizations and experts likely to be interested in the work of the SDMX projects and any other SDMX initiatives in order to foster their involvement

- Invite global participation wherever possible

- Make use of tele- and video-conferencing when appropriate