CCSA Conference on Data Quality for International Organizations

(Wiesbaden, May 27-28, 2004)

Data and Metadata Collection and Dissemination in the IMF Statistical Information System

Statistics Department
International Monetary Fund

Agenda

DATA

- Recent Developments in IT
- Recent Developments in Policies
- Tools to Support Policies
- Challenges Ahead

METADATA

- IMF Metadata Environment
- Metadata Management Issues
- Building Blocks for Improvement
- SDMX Initiative

Recent Developments

- Information Technology
  - Data management
    - new **Economic Data Facility (EDF)** for managing time series data
  - Data collection
    - new **Integrated Correspondents System (ICS)** for secure web-based electronic data collection
    - SDMX standards, e.g., **GESMES/TS** in progress
  - Publications
    - in-house production capabilities, web products, enhanced CD-ROM products (coming)

Recent Developments

- Policies and Procedures
  - New vision, goals and objectives for STA data management and publications
  - Examining optimal organization and staffing for STA data management and publications
  - Revising STA data management policies

Recent Developments

- Policies and Procedures (2)
  - Re-engineering STA work processes, adopting a common model across domains that support vision, goals and objectives and new organization and staffing
  - New training program to support change management, supported by communications strategy

Tools to Support Policies

- ICS for data collection brings automation and provides for more accurate and timely data
  - Excel and standard formats, e.g., GESMES/TS, SDMX-ML (in progress)
- EDF data management system has more validation tools to ensure quality
  - Numeric checks, cross-topic consistency checks, customized business rules for validation
- EDF supports metadata collection for time series at multiple levels to aid transparency
  - Footnotes and flags by data observation, time series, data set, country, topic, etc.

Challenges Ahead

- Limited data harmonization across countries (a constraint for automation)
  - Introducing standardized “report forms” for monetary statistics
  - Work towards content standards under SDMX umbrella (e.g., common code lists, standard key families, etc.)
- Technology varies widely across IMF membership
  - Countries that have already invested in GESMES/TS and XML technologies can build on these investments using SDMX standards
  - Countries that are just beginning to invest have an easy “entry point” that can serve multiple purposes (collection and dissemination) and partners
- Limited availability of “structural” metadata (time series and observation-level)
  - SDMX information model facilitates electronic exchange of “structural” metadata

IMF Metadata Environment

- More than two-thirds of the IMF membership disseminate information using the SDDS/GDDS framework and the IMF's Dissemination Standards Bulletin Board (DSBB)
  - Special Data Dissemination Standard (SDDS)—subscription has grown to 57 countries
  - General Data Dissemination System (GDDS)—metadata posted for 68 participants
- Information on 18 macroeconomic "*data categories*" (plus population, health, education, and poverty) is available using the SDDS/GDDS framework

IMF Metadata Environment

- The database supporting the IMF metadata repositories provides for
  - Dynamic generation of web pages
  - Query facilities providing information on detailed metadata content
  - Comparison of information on cross-country practices
  - Using Word templates to facilitate the updating of the metadata

Metadata Management Issues

- Collection
  - Metadata providers have no direct updating facilities (No metadata equivalent to the ICS)
  - Metadata maintenance remains resource intensive at both ends (the providers and the IMF)
  - Quality of content could be improved by a common vocabulary
- Dissemination
  - Limited flexibility in querying the metadata and using the outputs from the website

Building Blocks for Improvement

- SDDS
  - Recognized standard for the dissemination of “reference” metadata
- IMF’s Data Quality Assessment Framework
  - DQAF—Cascading structure to identify assessing quality
  - Rooted in UN *Fundamental Principles of Official Statistics*
  - SDDS/GDDS framework supports linkages to the data quality concepts of the DQAF
- SDMX initiative
  - Taking advantage of IT developments to build a model and the supporting tools for the open exchange of “reference” metadata

SDMX Initiative

For “reference” metadata exchange

- Develop a generic metadata model that will
  - Support exchange for a range of statistical products
  - Comprise the basic concepts and their data category specific extensions, established jointly with relevant authoritative bodies
    - SDDS application of the model—with flexible extensions for other applications
  - Allow for linkages between data and metadata
  - Provide the tools to produce and “read” XML files

SDMX Initiative

- Medium-term objectives—metadata exchange
  - Reduced cost in collecting and maintaining metadata (available in standard formats)
  - Increased quality (harmonization of concepts and terms for each data category)
  - Increased access (using XML, web services, and other tools)
  - Linkages to the data (for SDMX compliant organizations)