

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

Seminar on Statistical Data Collection
(Geneva, Switzerland, 25-27 September 2013)

Topic (i): Centralising data collection

**PROCESS AND CONTENT STANDARDISATION FOR DATA
COLLECTION**

Working Paper

Prepared by Olga Laveda and Eric Dery, Statistics Department, International Monetary Fund

I. Introduction

This paper describes the work the International Monetary Fund (IMF) is doing with Eurostat and the African Development Bank (AfDB) in standardising and streamlining all agencies' data collection from countries for their internal operational work and policy making. With the last financial crisis, it was clear that more detailed data and timely availability of data are critical for in-time response and better understanding/control of the crisis. The crisis also exposed data gaps in the existing data that were necessary for a better crisis management. To address this data issue, the IMF Statistics Department (STA) is working in multiple areas with its partners to streamline the data collection without increasing the reporting burden on the countries. Process and content standardisation are critical for achieving the goal. For this purpose, IMF is working with Eurostat and the Inter-Agency Task Force on International Data Sharing to collect on data using SDMX and with AfDB to leverage their OpenData Platform to streamline the country data collection of IMF datasets.

II. Business Context

In the wake of the 2008 financial crisis, the international community identified important data gaps that needed to be filled in order to create a global information system suitable for monitoring global financial and nonfinancial flows and positions comprehensively, as well as monitoring the ongoing impact of inter-connectedness. In particular, the report *The Financial Crisis and Information Gaps* prepared by the staff of the Financial Stability Board (FSB) Secretariat and the IMF and endorsed by the Group of Twenty (G-20) Finance Ministers and Central Bank Governors in November 2009, identified areas of data collection and processing that would need to be addressed by the IMF and other international organizations. It is anticipated that, primarily in response to the G-20 Data Gaps Initiative, the volume of data processed by STA would increase by multiple factors over the next few years. The increased demand is expected to have to be accommodated within existing budgetary resources and by minimising the increase in effort in reporting by the countries.

III. Objectives

STA plans to streamline data collection processes and respond to the impending business context together with a suggested approach for achieving these objectives.

Streamline data collection from African countries

Standardise content and process for data collection in collaboration with AfDB leveraging the Africa OpenData leading to reduction in reporting burden for countries.

Streamline national accounts data collection from all countries

Under the auspices of the Inter-Agency Group on Economic and Financial Statistics (IAG), the Bank for International Settlements (BIS), the European Central Bank (ECB), the Statistical Office of the European Communities (Eurostat), the IMF, the Organisation for Economic Co-operation and Development (OECD), the United Nations (UN) and the World Bank (WB) have agreed on general principles governing the data sharing of official statistics. The IAG mandates an international task force to develop and carry out two pilot projects for testing practical arrangements

IV. Business Implementation

A. AfDB OpenData Platform

1. Data Collection

As a part of “Standardise, Streamline and Automate” initiative, STA is reviewing its existing data flows and processes. Through this review exercise, STA is exploring options to improve and streamline their operations in order to meet the growing demand for data. One of the objectives is to improve timeliness and streamline data collection from countries.

2. Standardise collection content for comparison

In addition to the existing modes of data collection using excel, it is proposed to standardise collection content through the SDMX format. Data providers will be encouraged to move up the automation chain by disseminating data to the IMF through unified codes and standardised SDMX dimension and attributes across domains. This will greatly reduce the reporting burden on the country side and help in content comparisons.

- Standard / uniform collection structures (dimensions and attributes) for all collection domains / datasets
- Constraints at dataflow level for ease of reporting by dataset or report form
- Use of standard/unified codes for the codelists to enable collection automation and data comparisons

3. Data submission process through OpenData Portal

AfDB is implementing OpenData platform for all countries in the African region. IMF has partnered with AFDB to leverage their OpenData platform and collect data from the countries using SDMX. The future data submission process would include the following steps:

- Country uploads data once to their OpenData Platforms
- OpenData Platform
 - Maps data to standard codes
 - Enables data exchange using SDMX standards
- Country correspondents inform agencies to download data from the OpenData Platform

4. Impact on Reporting Countries / Authorities

Currently data provision from country authorities are requested by various organizations and data are in different formats and timings. Through this initiative the countries would update their data at one place and collection agencies would be able to pick from there. This would possibly result in:

Reduction in reporting burden for the countries.

Implementation of standardised codes

Enable machine to machine exchange of data using SDMX

B. IAG task force on international data sharing.

The IAG mandates an international task force to develop and carry out two pilot projects for testing practical arrangements implementing the general principles for two distinct cases. The first test case will put focus on the detailed set of sector accounts for G-20 economies (recommendation 15 of the G20 Data Gaps Initiative), to identify large macroeconomic imbalances that would require coordinated global preventive or corrective policy actions to promote sustainable and balanced growth as requested by the G-20 leaders in November 2010.

The second test case would put emphasis on a much wider representation of countries and international organisations, covering the main aggregates (GDP and expenditure components) and population from the System of National Accounts. This would particularly help to reduce the incoherent information users currently find for a single country in different international databases for some very widely used indicators.

1. Principles

Common statistical concepts and definitions

- Over the past decades the I&SOs have made significant progress in developing international statistical standards for official statistics.
- The System of National Accounts (SNA2008), the Balance of Payments Manual and International Investment Position Manual 6th Edition (BPM6) and other related international statistical standards provide a solid methodological basis for compiling economic and financial statistics on the basis of aligned concepts and definitions.
- The implementation of the new more aligned statistical methodological standards is a crucial step to compile official statistics and indicators that are globally comparable, timely and sound.

2. Globally harmonised reporting templates for statistics

- The adherence to the international statistical standards needs to be promoted through complementary *reporting requirements* organised around globally harmonised templates.
- These *globally harmonised templates* will ensure that key statistics for global economic surveillance can be made available in a comparable format for all economies.
- The globally harmonised templates:
 - Should specify *the domains, statistics or variables selected*.
 - Should be designed following a ‘modular’ approach, with a so-called ‘Tier 1 module’ for the core set of main variables and a ‘Tier 2 module’ for detailed breakdowns and a further ‘Tier 3 module’ referring to specificities of the reference economy.
 - Should specify *the frequency and timeliness* of the data to be reported (e.g. monthly/quarterly with a timeliness of T+30/60/90/180 days or annual data with a timeliness of T+3/6/9 months).
 - Should promote harmonized reporting templates based on *SDMX standard formats and codifications (DSDs)* to enable a smooth access and efficient sharing of template-based data between I&SOs¹.

3. Global data sharing of statistics

- The global data sharing arrangement aims at defining an *efficient exchange between the International and Supranational Organisations (I&SOs) of statistics* to govern the data flows from national authorities to the I&SOs, and also between the I&SOs.
- This efficient data exchange establishes a clear *allocation of responsibilities among I&SOs on data validation and dissemination* of the various sets of statistics.

¹ SDMX standard formats and codifications (DSDs) for Balance of Payments statistics and National Accounts will be introduced in 2013 and 2014, respectively.

- The data sharing should ensure that the economic and financial statistics displayed in the databases of all collaborating I&SOs are *identical* for similar statistical concepts (i.e. no vintage or methodological differences) and of the *highest quality* at any time.
- It goes without saying that the *national authorities* will be ultimately responsible for the quality of their data. If national data, as delivered to I&SOs need to be adjusted, this will only be done in close cooperation with the relevant national authorities.

4. Single point for data quality management

- For every domain of statistics required through harmonised reporting templates, one I&SO will be selected as the '*primary (validator and) disseminator*' of the data of a certain country.
- The selection of the relevant I&SOs will largely *take advantage of data sharing arrangements already in place*. This often is based on regional or supranational arrangements, which could also (but not necessarily) be reflected in legal acts.

5. Scope of the task force

- The task force is organising two pilot tests for the implementation of the general principles on data sharing between international and supranational organisations.
- Pilot 1 has the objective to test the data sharing in the area of institutional sector accounts within the scope of the G20 (recommendation 15 of the G20 Data Gaps Initiative).
- Pilot 2 has the objective to test the data sharing for a very basic data set of main aggregates and population data with the widest possible representation of countries and organisations.
- If more appropriate, the task force may decide after a kick-off to work in two sub-groups on those two pilots, however closely coordinating common elements.

6. Technical Modalities of Data Exchange

The pilot exercise will encompass the transmission of data messages with regards to the agreed template for GDP main aggregates and population statistics; the data transmission will make use of SDMX-ML messages using the now nearly finalized DSD for national accounts.

The scope of the data transmission will encompass a list of countries for which it is established that quarterly main aggregates are regularly produced and disseminated. This list, it has been established, is a super set of the data currently collected by the OECD and Eurostat, which implies that for several countries a primary validator/disseminator IO needs to be assigned.

An automated data exchange through web service and data API will be implemented with a registry service and sandbox. The Sandbox should support the web service call for extracting the data instead of manual downloading of SDMX message files.