

Distr.
GENERAL

Working Paper
27 February 2014

ENGLISH ONLY

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE (ECE)
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN UNION (EUROSTAT)**

**ORGANISATION FOR ECONOMIC COOPERATION
AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

**UNITED NATIONS
ECONOMIC AND SOCIAL COMMISSION
FOR ASIA AND THE PACIFIC (ESCAP)**

Meeting on the Management of Statistical Information Systems (MSIS 2014)
(Dublin, Ireland and Manila, Philippines 14-16 April 2014)

Topic (ii): Standards-based modernization

Implementation of SDMX Standards and Guidelines: A Pilot Project in Asia and the Pacific

Prepared by:

Denis Ward, Consultant, Asian Development Bank (ADB)

Artur Andrysiak, Asian Development Bank (ADB)

Eric Hermouet, Economic and Social Commission for Asia and the Pacific (ESCAP)

Marko Javorsek, Economic and Social Commission for Asia and the Pacific (ESCAP)

I. Introduction

The purpose of this paper is twofold. First, to outline the joint Asian Development Bank (ADB) and United Nations Economic and Social Commission for Asia and the Pacific (ESCAP) pilot project to develop the capacity of countries in Asia and the Pacific to implement SDMX standards and guidelines. Second, to highlight some of the issues confronting an international agency acquiring knowledge of, and embarking on, the implementation of SDMX standards for developing countries.

The paper outlines:

- the objectives of the initiative (Section II);
- the results of recent investigations to identify existing SDMX implementation tools (Section III) developed by the SDMX sponsoring agencies that could be readily utilised by countries in Asia and the Pacific. This Section also lists current SDMX implementations identified to date that utilise some of these tools;
- recent work on the development of Data Structure Definitions (DSDs) and Metadata Structure Definitions (MSDs) – Section IV.

The paper concludes with a brief reflexion on SDMX implementations in the light of recent investigations for the initiative (in Section V), and with an outline of future plans / tasks for the joint ADB – ESCAP initiative (Section VI).

Project implementation for the joint ADB – ESCAP initiative initially entailed obtaining information about current SDMX Technical Standards and Content Oriented Guidelines. This was followed by investigation of implementation tools that have been developed by international organisations and national agencies to identify those that are most relevant for countries in Asia and the Pacific. This work has been undertaken concurrently with discussions with countries in the region to gauge their knowledge of SDMX and to identify their aims, objectives and capacity development needs with regards to SDMX implementation.

Information obtained from these investigations is summarised below. It should be emphasised that much of the information provided below on SDMX implementations is partial and incomplete and that the process of gathering details from international agencies and NSOs is still in progress.

II. Project Objectives

The overall objective of the joint ADB - ESCAP initiative is for statistical agencies in the region to work together improve the efficiency of data and metadata exchange between national statistical agencies and international organizations through the ongoing use of SDMX standards and guidelines. At moment these standards are used only by a small number of national statistical agencies in the region.

The project aims to build the capacity of developing countries in the region to apply SDMX standards and guidelines through the development of required SDMX DSDs and MSDs or the application of existing DSDs / MSDs developed on global level or for other specific projects.

Beyond the immediate result of reducing the data processing burden on both senders and recipients of the data and metadata, countries in the region will benefit from gaining a practical knowledge of SDMX standards and implementation tools. This will serve as a basis for national agencies to further develop their SDMX data dissemination infrastructure with other organisations across a range of statistical domains, and as such decrease the overall response burden on the NSS. At a regional and global level the immediate benefit will include enhanced common data management and improved quality of data reporting through better coordinated data and metadata exchanges.

More specifically the project will focus on:

- The adaptation and use of DSDs and MSDs for economic data and their related metadata. In the initial stage the use of DSDs / MSDs will be limited to the statistical indicators disseminated by the ADB in their annual publication, *Key Indicators for Asia and the Pacific* (KI) and related data products¹. This will take into account the work of other international organizations collecting economic data, such as the IMF, UN, World Bank, OECD, Eurostat, etc., in order to maximise the reusability of components developed under the project.

In the pilot phase of this project, proof of concept will focus on the exchange of data and metadata for national accounts and balance of payments for which global DSDs have recently been developed by the SDMX sponsoring agencies. At a later stage, ADB and ESCAP might work with other agencies to identify other relevant (and hopefully global) DSDs for other statistical domains through multilateral forums such as the SDMX Global Conference; SDMX Expert Group; and the Management of Statistical Information Systems (MSIS), etc.

¹ The Key Indicators publication currently disseminates around 180 annual series for individual countries in the Asia and Pacific region. These include series in the following domains: national accounts; balance of payments; production; energy; prices; labour force; money and banking; government finance; external trade; international reserves; exchange rates; external indebtedness; and population. Over 120 series for national accounts and 20 BoP series are disseminated. Refer link to *Key Indicators for Asia and the Pacific 2013* - <http://www.adb.org/publications/key-indicators-asia-and-pacific-2013>

- Using (or adapting) existing SDMX implementation tools (including web services, national SDMX data repository tools, mapping tools, etc) that have already been developed by the SDMX sponsoring agencies – refer Section III below.
 - Working initially with a small number of pilot countries in the Asia and Pacific region to identify the relevant implementation tools, barriers to national implementation, etc. The pilot countries are: Malaysia and Thailand, as well as Australia and New Zealand (with the latter two playing the roles of technical advisors, as they have already implemented the SDMX tools).. The project will, however, aim at being replicable in other Asia-Pacific countries beyond the initial pilot countries. Criteria for the selection of existing implementation tools include:
 - ✓ availability of a set or suite interoperable tools including: web services, SDMX data repository, mapping;
 - ✓ ability of tools to link to a heterogeneous range of national database environments, in particular, those in developing and small-sized NSSs;
 - ✓ availability of both technical and non-technical documentation outlining functionality and use of the tools. Documentation should also outline the in detail the prerequisites for the use of mapping tool;
 - ✓ availability of training courses and online tutorials.

III. SDMX Implementation Tools and Implementations

The following information was gleaned from papers and presentations given at meetings of the SDMX Global Conference and Expert Group meetings, the SDMX website, websites of SDMX sponsoring agencies and recent discussions with OECD, Eurostat, IMF, AfDB and NSOs. The information gathering process is still in progress.

Over the last five years or so the SDMX sponsoring agencies, other international organisations and national statistical agencies have developed a range of tools that incorporate the SDMX Technical Standards and Content-Oriented Guidelines (COGs). These tools have been developed to facilitate a number of specific implementation initiatives undertaken by each agency.

There currently appears to be two main sets of SDMX implementation tools that the ADB-ESCAP SDMX initiative could use or adapt, though their actual relevance for some of the less developed countries in the Asia and Pacific region is yet to be firmed up². The implementation tools comprise:

- Eurostat's SDMX-RI which has been used extensively by EU countries for the exchange of data and metadata. The OECD appears to be going in the direction of adopting this implementation.
- IMF's SDDS Plus which has used SDMX for exchange of the National Summary Data Page (NSDP). It has also been adopted by the African Development Bank (AfDB) for their Open Data Platform. The use of the tools by developing countries in Africa could be relevant for those in the Asia and Pacific region.

At the moment it is not clear as to how interoperable these implementations are, and whether it would be possible to use a combination of some of the tools developed. Theoretically, the inclusion of the COGs in the tools should make them interoperable, though some papers recently presented at SDMX global conferences and expert meetings have questioned this. Also, some of the DSDs prepared for these implementations appear to

² The different tools, demonstrations on their use, etc., are available in the SDMX Tools Repository located on the SDMX website at <http://www.oecd.org/std/47610077.pdf>.

duplicate and overlap. This could mean that national agencies providing data for each implementation would require a separate mapping.

Information on these implementation tools is still being gathered and the following details will be expanded / revised.

A. Eurostat’s SDMX Reference Infrastructure (SDMX-RI)

Eurostat’s universal framework for modern data provision, SDMX-RI, provides a set of building blocks allowing a statistical office to expose data to the external world based on access rights. The building blocks are designed to provide data and structural metadata based on mappings to each organization's dissemination data warehouse. The infrastructure uses SDMX standards and was developed to simplify the exchange of data and metadata. It provides standard software and components, allowing individual statistical organisations to interact and exchange their data using the same software and methodology. SDMX-RI’s modular approach enables users to use either part or the entire infrastructure.

The SDMX-RI building blocks comprise: web service; web client; mapping assistant; test client³.

Eurostat has also developed a number of training videos, work books and self-assessment resources to assist the implementation of the tools in their SDMX-RI.

In addition, Eurostat has engaged the services of a consulting firm to prepare training packages. Eurostat has already helped present SDMX in developing countries in Latin America, in partnership with ECLAC. Eurostat already offers regular training on SDMX in Luxembourg and Europe and could also send a trainer to the Asia-Pacific region.

B. IMF SDDS Plus⁴

SDDS Plus⁵ is a new tier in the IMF’s Data Standards Initiatives. SDMX is the dissemination format selected for SDDS Plus, replacing HTML. Adherence to SDDS Plus is open to all SDDS subscribers, especially those with systemically important financial sectors. SDDS Plus has nine additional data categories to the SDDS and longer time series – requirement is for five years of data for all prescribed components.

The National Summary Data Page (NSDP) in HTML is replaced by a series of SDMX-ML files or an SDMX web service call (one per data category).

For countries without an existing SDMX infrastructure, the IMF provides an Excel to SDMX converter specific for the NSDP-DSD.

In addition to the NSDP-DSD, IMF will support any “DSD for Global Use” developed under the auspices of SDMX Sponsors. Countries are free to decide which DSD (NSDP-DSD; Global DSDs) they wish to use for each data category.

The AfDB’s Open Data Platform is adopting SDMX and the NSDP-DSD for the dissemination of economic and financial statistics – refer below.

³ Source: Eurostat

⁴ Source: IMF

⁵ Overview of SDDS Plus: <http://www.oecd.org/std/SDMX%202013%20Session%203.11%20-%20The%20benefits%20of%20SDMX%20for%20SDDS%20Plus.pdf>

C. IMF – AfDB Open Data Platform

The ODP is a cloud-based platform. Its development was sponsored by the African Development Bank (AfDB) who makes it available to data producing agencies in Africa. It does not require any infrastructure investment on the part of the using agency.

The IMF is setting up each relevant agency in a country (NSO, central bank, etc) with its own ODP, which it fully governs (the Fund provides a suggested governance model). Data can be loaded into the ODP via Excel or CSV. The data can be loaded using national codes and the mapping to SDMX codes (the EcoFin DSD, the same used for SDDS Plus) takes place automatically in the platform. The data then become available to users via the user interface and supports machine-to-machine exchange using the ODP SDMX web service.

The data can also be downloaded in various formats (Excel, CSV, pdf, SDMX) via the user interface. The IMF has already started using the ODP SDMX web service as its data transmission (pull mode) system for ODP-using countries.

The EcoFin DSD is suitable for economic and financial statistics, but other DSD(s) would be needed for socio-demographic data. The IMF has started working with ILO and FAO (and AfDB) in that direction⁶.

The AfDB reported two main challenges in the implementation of the ODP across countries in the region.

- To have countries upload and update their datasets into the ODP.
- Each country has its own indicators, definitions, classifications, which may depart from international standards. Hence the importance of mapping with the international DSDs. The AfDB has started this work with the IMF, where the data are being uploaded in the ODP, mapped to the IMF codes and DSDs for macrodata. It is still being piloted in three countries and will be extended to others once the process has been finalized⁷.

The following table lists a number of SDMX implementations that have been identified to date using either the tools described above or other SDMX tools. These initiatives have been conducted by the sponsoring agencies and other international agencies in concert with a number of NSOs and central banks. The initiatives utilise SDMX technical standards and guidelines to facilitate:

- data collection (using either push or pull modes);
- data dissemination; or
- data warehousing.

Table 1: SDMX implementation projects identified to date [to be expanded / revised]

Agency(s) involved	Summary
OECD, UNSD, INEGI	To replace Excel files by SDMX flows for: short Term Economic Indicators; merchandise trade statistics; quarterly GDP at constant prices; MDGs.
OECD	SDMX data exchange pilot program using short-term economic statistics (STES). OECD Member countries were invited to participate in this exercise in 2011 and more than 20 countries are now participating in this exercise. Project aim is to improve the STES data collection process. Currently at various stages of implementation with several countries.
Statistics Korea, IBGE	Both NSOs have projects to transmit STES data via SDMX. Both agencies are using this exercise as a catalyst to promote further such exchanges throughout their organisations and with other agencies.

⁶ Source: IMF

⁷ Sources: IMF and AfDB

Agency(s) involved	Summary
Brazil	
UNSD, DevInfo (MDG project)	<p>Both organisations have been working together to deploy a system whereby development data is transmitted in SDMX. The MDG project gives the participants their first SDMX experience with well-known and easy-to-use platform and tools, such as DevInfo software or their own existing dissemination systems.</p> <p>MDG Metadata Structure Definition, available as a draft, was adapted and used by the project for the sharing of reference metadata.</p> <p>SDMX exchange established, for the most part, between international organizations or advanced countries and international organizations.</p>
World Bank, Commonwealth Secretariat (ComSec), External Debt Data	Collecting and Maintaining External Debt Data - Debtor Reporting System (DRS) ⁸ . World Bank and ComSec worked on adopting SDMX standards. System piloted with data of three countries
Bank of Italy	SDMX and Web Services provide the means for strongly simplifying data retrieval from external sources. Reduce the amount of external data that an organization needs to replicate in the internal systems, without impacting data process efficiency.
INEGI, Mexico	Uses SDMX via a number of initiatives in data collection, publication and dissemination.
Rosstat, Russia	Implementation of the SDMX standard by Rosstat to help integration in a decentralised system. 64 state bodies are making data available as SDMX via a single portal (UNISIS).
ILO	ILOStat development in 2011 uses SDMX interface for data collection and dissemination.
Swiss FSO	Use of SDMX for their statistical metadata system.
Australian Bureau of Statistics	

IV. SDMX Data Structure Definitions and Metadata Structure Definitions

A. Data Structure Definitions (DSDs)

The sponsoring agencies and other international organisations have developed DSDs for some of the statistical domains described in the COG's Statistical Subject Matter Domains. These include the priority national accounts and BoP domains for the ADB-ESCAP SDMX initiative. Listed below are the DSDs that have been prepared to date or are in the process of being developed, together with the name of the maintenance agency.

A distinction can be drawn between DSDs that are global in nature, whose intended use extends beyond the corporate requirements of the maintenance agency, and those that are primarily intended to meet a specific corporate need. The former DSD's have been developed after extensive consultation between international and national agencies whilst the level of consultation for the latter is more truncated. Ideally, DSDs should be developed that support the reporting needs of a number of international agencies. Examples of global DSDs include those developed for national accounts and BoP. An example of a "corporate" DSD includes the one

⁸ Description of project: <http://www.oecd.org/std/SDMX%202013%20Session%203.3%20-%20Debt%20Reporting%20by%20Developing%20Countries.pdf>

developed for the OECD's short-term economic statistics (STES). Some of the indicators included in the STES DSD duplicate those in the two global DSDs referred to above.

Table 2: DSDs that have been developed or are in the process of being developed [to be expanded / revised]

DSD	Maintenance agency	Nature of DSD
Balance of Payments (BOP)	IMF	global
National Accounts (NA) [30 September 2013]	Eurostat	global
Foreign Direct Investment (FDI)	OECD	
Government Finance Statistics (GFS)		
Eco-Fin	IMF	corporate
MDGs		global
Debt Reporting by Developing Countries		
R&D data	OECD, UNESCO	
National Statistics Data Page for SDDS Plus	IMF	corporate
Fisheries catch statistics	FAO	
Fisheries landings statistics	Eurostat	
Waste statistics		
Education	Eurostat	
Health	Eurostat	
Air transport statistics		
Pesticide statistics		
International merchandise trade statistics		
Job vacancy statistics		
Labour cost index		
Short-term Economic Indicators	OECD	corporate

Because of the need to assess the quality of the indicators disseminated by countries across the Asia and Pacific region, in particular, by those in developing countries, the structuring and efficient exchange of metadata is a key element of the joint ADB – ESCAP SDMX initiative.

B. Metadata Structure Definitions (MSDs)

To date, the development of MSDs and implementation of metadata exchange has been undertaken by the sponsoring agencies in the context of corporate initiatives such as Eurostat's ESS Metadata Handler, IMF's Data Quality Assessment Framework for the SDDS and GDDS; and UNSD's MDGs. Although the metadata items included these initiatives varies there is considerable overlap and each are based on the same metadata concepts and definitions.

Table 3: Metadata exchange initiatives [to be expanded]

Agency	Summary
Eurostat's ESS Metadata Handler (ESS MH)	<p>Eurostat's ESS MH comprises IT applications for managing the production, exchange and dissemination of European statistical system metadata. It accommodates the Euro-SDMX Metadata Structure (ESMS) and the ESS Standard for Quality Report Structure (ESQRS) for the production, exchange and dissemination of the national and Eurostat metadata files.</p> <p>The ESS MH stimulates the use of the SDMX based ESS metadata standards within the ESS and the further integration of the ESS metadata systems. This enables the on-line production of the national ESMS and ESQRS files and their exchange within the ESS.</p> <p>The ESMS facilitates the documentation of methodologies, quality and the</p>

Agency	Summary
	<p>statistical production processes in general. It uses 21 high-level concepts, with a limited breakdown of sub-items, strictly derived from the list of cross domain concepts in the SDMX Content Oriented Guidelines – refer Annex 2 for a listing of the Euro-SDMX Metadata Structure concepts. When fully implemented, there will be more than 2000 Eurostat and national ESMS files produced and disseminated. Eurostat⁹.</p> <p>The ESQRS provides the concept base for the ongoing and systematic monitoring and reporting of the quality of statistics produced within the European Statistical System. Statistics. This will enable producers to monitor the quality of the statistics by concentrating on the main quality concepts, namely: relevance; accuracy; timeliness and punctuality; accessibility and clarity; comparability; and coherence.</p> <p>The ESQRS uses eight main concepts taken from the SDMX cross-domain concepts plus more detailed sub-concepts measuring data quality. It is a metadata structure which is used across all statistical domains.</p> <p>[Source: Eurostat]</p>
IMF Data Quality Assessment Framework (DQAF) / AfDB ODP	<p>Structured metadata on the basis of templates covered are provided by around 180 countries subscribing to the IMF’s SDDS and GDDS that facilitate assessments of data quality across a range of statistical domains. The metadata concepts used in the templates are by and large consistent with those used by Eurostat, OECD, etc.</p> <p>At the moment, the ODP does not include a metadata component which is something the organisations will have to look into at a later phase. Immediate plans are to provide countries with a (basically free) data dissemination platform that meet the “Open Data” criteria.</p> <p>[Source: IMF]</p>
UNSD’s MDGs	

V. Reflection on SDMX Issues

Recent investigations on SDMX standards and guidelines and implementations by international and national statistical agencies for the joint ADB – ESCAP initiative have brought to light a number of issues which could be discussed with the four pilot countries and at the MSIS meetings to be held in April 2014. These issues have been raised in the past at SDMX Global Conferences and Expert Group meetings and some are currently being addressed by the SDMX Working Groups, etc.

A. Need for Global SDMX Registry

The need for a Global SDMX Registry has been recognised for some time and such a registry is expected to be available in 2014. The amount of effort required first to identify DSDs, MSDs, SDMX implementations and then assess their relevance for the ADB – ESCAP initiative (in particular, for developing countries in the Asia and Pacific regions further highlights the need for the Global Registry.

If possible, such a registry should be well structured and extend beyond the mere listing of SDMX artifacts and include where possible how some of these artifacts relate to each other. For example, are they interoperable, do they duplicate each other (e.g. global v. “corporate” DSDs)?

B. Audience for SDMX information on SDMX artifacts

⁹ Refer <http://www.oecd.org/std/47624538.pdf>

Reading through papers prepared and presented at SDMX conferences and meetings and SDMX material on the websites of national and international agencies, one gets the impression that much of the information is aimed at IT and / or information dissemination specialists. Some papers presented over the last five years have commented that much SDMX material / documentation is quite technical and have pointed to the need for it to suit a broader audience and to also include information aimed at the statistical domain specialists, and for statisticians and management in developing countries. For example:

- It is too easy to form an impression that the application of SDMX standards and guidelines are only relevant for developed national statistical systems and are of little direct benefit, or too costly to implement, either in developing countries or in small-sized statistical systems.
- More specific information is required on mapping national statistical concepts to SDMX concepts and codes. Mapping tools have been developed by some of the sponsoring agencies, however more detailed information is needed on what information is required from statistical subject matter specialists in national statistical agencies contemplating the adoption of SDMX guidelines for data / metadata exchange, data dissemination, etc. For example, information on national data and metadata infrastructures, conformity of national statistical series with SDMX concepts, data reported to various international organisations, etc. Also, what happens where national concepts are inconsistent with SDMX concepts?

C. Updating of SDMX website

Much of the information currently on the SDMX website is out of date and needs to be revised, in particular, information on: SDMX activities across various statistical subject matter domains; and information on SDMX implementations developed by sponsoring agencies, other international organisations and national statistical agencies. Information on implementations currently on the website appears not to have been updated since 2007. Information accessible via the SDMX Tools Repository is more recent though entries are not time stamped and it doesn't appear to cover activities by all national and international agencies.

D. Current and future SDMX development activities

There is also a need for a brief, readily accessible outline of current and future work by the SDMX Working Groups and relevant activities by each of the sponsoring agencies on implementations, etc. Some information is available in papers presented at conferences but it appears to be partial and is difficult to locate. Furthermore, it is sometimes difficult to determine whether an implementation is still being developed, at the pilot stage of implementation, or fully operational.

E. Review of existing SDMX tools

There is a need to undertake a comparative review of the existing SDMX tools developed by various organisations at the global level. As mentioned in Section II, various organisations and private entities have developed SDMX tools. Each of these tools have various advantages and disadvantages for certain countries or environments. A global group, such as one of the Modernization Committees functioning under HLG, could undertake such a comparative review. Such an independent review would be extremely useful to any country or organisation starting to use SDMX, such as the initiative under this project.

VI. Future Tasks / Plans for the Joint ADB – ESCAP SDMX Initiative

1. Discussions with key statistical agencies (NSOs, central banks) in countries to ascertain:
 - What they want to get out of their involvement in the joint ADB – ESCAP initiative? What are their objectives? Are they interested in using SDMX for data collection; data dissemination and reporting; or data warehousing? Their current knowledge of SDMX standards and guidelines and existing implementation tools.

- Information on aspects of the statistical institutional environment relevant to the proposed pilot study, for example: agency(ies) responsible for the compilation of national accounts, BoP and other KI economic statistics; conformity of these indicators to international statistical standards and series collected by international organisations such as IMF, UN, World Bank, OECD, etc; database environment(s); who would be the main drivers for SDMX implementation within national agencies – IT, dissemination and marketing, subject matter areas; what data are currently reported to international organisations, who does this and how are data transmitted?
 - Implementation barriers – both national and in respect to the SDMX standards and existing implementations – technical; knowledge and skill base; ease of accessing and understanding SDMX documentation; resource issues.
2. Identification of appropriate SDMX implementation tools, etc., to utilise in the implementation of the ADB – ESCAP initiative that suit the requirements / environments of countries in the Asia and Pacific region, initially in the pilot countries. This phase will involve further discussions with IMF, Eurostat, etc., to:
- As a prelude to examining the coverage of and utilisation of existing DSDs / MSDs, identify national account and BoP statistical series currently collected by those agencies and to match them against series disseminated in ADB’s Key Indicators publication. For example, from information obtained from the IMF we understand that there is a fundamental difference between the EcoFin and BoP DSDs. The former is designed for web-dissemination / mobile apps / and easy access to data; the latter is meant to respond to very detailed reporting requirements and can only work in a reporting mode. The IMF has plans to provide a mapping between the two and will work on this as soon as the SDMX Global Registry becomes available. Also, in the longer term other statistical domains in the ADB publication will also need to be matched.

As mentioned previously, in order to avoid duplication of effort at both national and international levels the ADB-ESCAP initiative will give priority to the use of global DSDs where they exist, followed by existing corporate DSDs. Additional DSDs will only be developed by the ADB-ESCAP initiative as a last resort.

- Assess availability and suitability of existing documentation on implementation tools as well as training required and training opportunities provided by Eurostat / IMF.
3. Prepare action plan and recommendations for other countries in Asia and the Pacific arising from experiences in pilot countries, particularly with respect to capacity development requirements to utilise common DSDs / MSDs and implementation tools to support automated data exchange.