"Science Stays True Here"
Journal of Mathematics and Statistical Science, 22-30 | Science Signpost Publishing

SDMX, a Key Standard for Central Banks' Statistics

Bruno Tissot[1]

BIS, Basel, Switzerland – Bruno.Tissot@bis.org

Abstract

Having structured and well-documented statistics on the economy and the financial system is an essential objective for central banks. This is reinforced by the fact that they collect, compile, share and disseminate a wide range of data – even more so since the Great Financial Crisis (GFC) of 2007/09 and the significant data needs unveiled on this occasion. The Statistical Data and Metadata eXchange (SDMX) standard provides both an information model and a syntax that enables standardisation for the structure as well as the presentation of statistical datasets and supporting metadata. Triggered and supported by international organisations, this standard is now widely used by central banks to facilitate and streamline their statistical work, especially as regards the IT-related aspects. This is indeed a key conclusion of a related survey published by the Irving Fisher Committee on Central Bank Statistics (IFC) in 2016. Yet the SDMX standard needs to evolve further, not least to better deal with the handling of large micro data sets that are in high demand after the crisis.

Keywords: central banking, great financial crisis, statistical system, micro data.

1. Introduction

In 2016, the IFC published a report on central banks' use of and interest in SDMX (Statistical Data and Metadata eXchange; cf IFC (2016)). SDMX is an ISO standard to describe statistical data and metadata, normalise their exchange and improve their efficient sharing. It provides an integrated approach to facilitating statistical data and metadata exchange, enabling interoperable implementations within and between systems concerned with the exchange, reporting and dissemination of statistical data and their related meta-information.[2]

The IFC report was derived from a broader online survey organised by the SDMX sponsor organisations. The aim of this general survey was to measure the acceptance and implementation of

[1] The views expressed are those of the author and do not necessarily reflect those of the BIS or the IFC. The text benefited from valuable comments from Heinrich Ehrmann and Xavier Sosnovsky.
[2] See https://sdmx.org/.

SDMX in the international community of official statistics (ie central banks, statistical offices, international organisations) in general. As regards the more specific IFC report, it mainly focused on three areas of importance to central banks: (i) their experience with the SDMX standards; (ii) their degree of satisfaction; and (iii) their expectations regarding potential developments in the standards.

2. The SDMX Standard: General Features

In 2001, the Bank for International Settlements (BIS), the European Central Bank (ECB), Eurostat, the International Monetary Fund (IMF), the Organisation for Economic Cooperation and Development (OECD), the World Bank and the United Nations Statistics Division (UNSD) – the seven sponsor organisations of SDMX – joined forces to develop more efficient processes for the exchange and sharing of statistical data and metadata. These seven sponsor organisations called this initiative Statistical Data and Metadata eXchange (SDMX). The objective was to give easy access to statistical data as well as to facilitate access to metadata that make the data more meaningful, comparable and usable.

In a nutshell, SDMX provides an integrated approach to:

- facilitating statistical data and metadata exchange:
- enabling interoperable implementations within and between systems concerned with the exchange of information; and
- processing and disseminating statistical data and their related meta-information.

The SDMX information model covers and describes the key concepts around statistical data, metadata and data exchange processes (SDMX SWG (2016)). This model comprises, among others, the following elements: descriptor concepts (ie concepts associated with the statistical data) as well as the types of these concepts (dimension, attribute or measure); the packaging structure (ie observation level, series level, data set level); the keys (grouping the various dimensions for a particular set of data); the code list (defining the possible values for a dimension). All this information is comprised in a specific Data Structure Definition (DSD, or "key family"), which specifies a set of concepts (eg dimensions, attributes, code lists – such artefacts are also referred to as "structural metadata") to describe and identify a set of data. Additional explanatory information is called "reference metadata", which describes the content, methodology and quality of the data. In SDMX terminology, the so-called Data and Metadata Structure Definitions – DSDs and MSDs – are made available in the SDMX Global Registry.[3]

[3] This is only the case for the Global DSDs and the SDMX cross-domain concepts and codelists. The other DSDs and MSDs maintained by the individual organisations (for instance by the BIS for the banking statistics, see below) are not available in the Global Registry.

The development of the SDMX information model has required important work on the technical specifications, as well as the provision of IT tools for data validation, conversion and dissemination. The related complexity has certainly raised a number of challenges (Cook and Soto (2015)). Yet the related technical standard is now widely available in an XML implementation format that allows an automated exchange and processing of statistical data. In 2013, SDMX was also published as an ISO International Standard (IS; 17369:2013)[4] to encourage and enable the widest possible use of the standard by organisations and companies involved with statistical data exchange.

This IT technical work has gone hand in hand with statistical work, ie the creation of internationally agreed structures and codelists for reporting/exchanging statistical data in a number of domains, eg national accounts, balance of payments and government finance.

One telling example is the reporting of the BIS international banking statistics (IBS).[5] Reporting authorities are encouraged to submit the IBS to the BIS using the SDMX standard. To this end, specific technical guidelines have been set up that provide practical information on how to submit data to the BIS. In particular, specific codes for submitting the banking statistics are provided in the data structure definitions. In addition, Excel templates are available as a visual aid for data structure, and to aid reporting authorities who are unable to create SDMX files directly from their system in reporting data to the BIS. Specific SDMX technical guidelines and data structures are made available to cover the locational banking statistics (LBS) by residence and by nationality as well as the consolidated banking statistics(CBS), with specific "mapping files" to organise both the reporting of data to the BIS and the subsequent dissemination of data by the BIS.[6]

SDMX now supports a wide range of international data exchange activities and is widely used around the world.[7] Many national and international modernisation initiatives use SDMX to improve and standardise data and metadata exchange and dissemination, to make data usage easier, to maintain statistical quality and to save production and dissemination costs. This progress has been acknowledged by the international community, for instance by the Data Gaps Initiative (DGI) endorsed by the G20.

---

4 See https://www.iso.org/standard/52500.html. ISO 17369:2013 is applicable to any organisation that has a need to manage the reporting, exchange and dissemination of its statistical data and related metadata. The information model at the core of ISO 17369:2013 has been developed to support statistics as collected and used by governmental and supra-national statistical organisations, and this model is also applicable to other organisational contexts involving statistical data and related metadata.

5 For an introduction to BIS statistics, see BIS (2015).

6 For the associated IBS reporting guidelines, see BIS (2013). The related SDMX technical guidelines and the DSDs are available on http://www.bis.org/statistics/bankstatsguide_tech.pdf and http://www.bis.org/statistics/dsd_lbs.pdf (for the LBS) and http://www.bis.org/statistics/dsd_cbs.pdf (for the CBS), respectively.

7 For an example of the analysis of the requirements related to data and metadata exchange between organisations, see for instance Denk and Grossmann (2011).

Indeed, the second phase of the DGI initiated in 2016 comprised a recommendation No 20 on International Data Cooperation and Communication which asks the Inter-Agency Group on Economic and Financial Statistics (which comprises the seven SDMX sponsor organisations) to "*foster improved international data cooperation among international organisations and support timely standardized transmission of data through internationally agreed formats (eg, SDMX), to reduce the burden on reporting economies, and promote outreach to users*" (IMF and FSB (2015)). In introducing this recommendation, it was in particular recommended to promote efficiencies in data supply including through (i) the promotion of the SDMX standard for the dissemination and sharing of official statistics, (ii) the creation of DSDs by statistical domain, and (iii) the reduction of overlap in transmissions of data to international organisations by national authorities.

In this context, the SDMX sponsors have outlined a series of strategic objectives in a *SDMX Roadmap for 2020*,[8] which are:

- strengthening the implementation of SDMX;
- making data usage easier via SDMX (especially for policy use);
- using SDMX to modernise statistical processes, as well as continuously improving the standards and IT infrastructure; and
- improving communication in general, including a better interaction between international partners.

3. Central Banks' Use of the SDMX Standard

According to their responses[9] to the IFC survey published in 2016, central banks' use of, and interest in, the SDMX standard can be summarised along the following main dimensions:

– **SDMX usage**. The use of SDMX is fairly common amongst central banks. Two thirds of respondents already use the standards, and more than one tenth of the group are planning to start using them soon. However, one in four of the respondents do not plan to use SDMX in the near future.

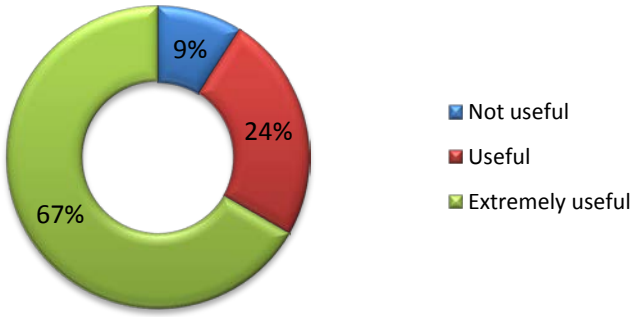
– **Satisfaction**. The vast majority of central banks using SDMX find the implementation "very useful" or "useful" (Graph 1). Less than one tenth of the respondents do not share this positive general view.

– **International data exchanges**. Most central banks already exchange or share data with international organisations using the SDMX standards, and in particular with the IMF and the BIS. In

8 See https://sdmx.org/?sdmx_news=sdmx-roadmap-2020.
9 Fifty-three central banks worldwide participated in this survey.

addition, European central banks also participate actively in data-sharing arrangements with the ECB and Eurostat.

How useful would you consider that the use of SDMX would be for your organisation? **Graph 1**



– **Statistical domains**. The use of SDMX standards to share data is still concentrated in a few statistical areas. Central banks use the SDMX standards mainly in a limited number of statistical areas for which international data structure definitions (DSDs) have been clearly established, eg balance of payments (BoP), sectoral national accounts (SNA), foreign direct investment and SDDS – the Special Data Dissemination Standard established by the International Monetary Fund to guide countries in the provision of their economic and financial data to the public.[10] The versions of the standards most commonly used are SDMX-ML 2.0 and SDMX-ML 2.1.

– **Data collection versus dissemination**. For most central banks, SDMX implementation is directed towards data dissemination (Graph 2). Data and metadata collection are comparatively less important objectives. From this perspective, it should be noted that for most central banks the concept of "dissemination" includes their reporting to international organisations.

– **Priorities**. Looking ahead, macroeconomic and finance statistics (in addition to the areas already

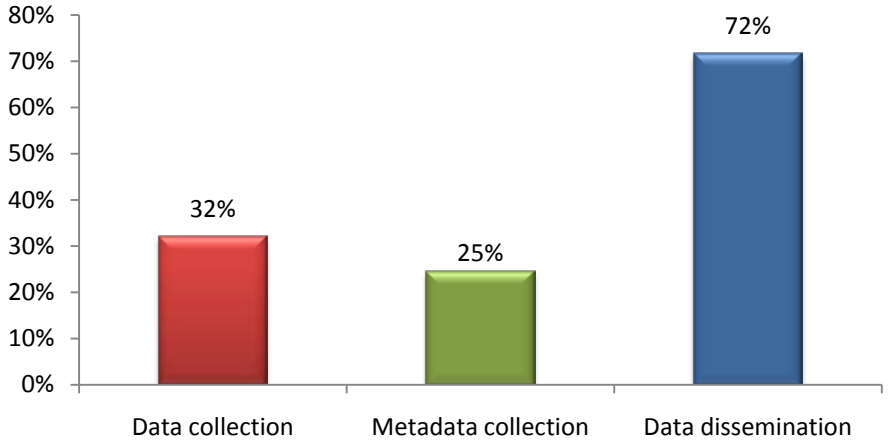
[10] The Special Data Dissemination Standard (SDDS) was established by the International Monetary Fund (IMF) to guide members that have, or might seek, access to international capital markets in the provision of their economic and financial data to the public. See http://dsbb.imf.org/Pages/SDDS/Overview.aspx. For the use of the SDMX as a platform supporting the SDDS, see IMF (2015).

covered such as SNA/BoP), international trade statistics, business statistics and prices are the four priority domains for developing new SDMX data structures.

– **Data reporting**. SDMX is very convenient for the providers of statistics. A key benefit of implementing the standards for central banks is to ease the burden of reporting to international organisations.

– **Challenges**. The main challenge is that internationally agreed data structure definitions (DSDs) and metadata structure definitions (MSDs) are still missing for some important statistical domains. The lack of adequate human resources and training capacities, as well as difficulties getting subject matter statisticians' support, are also widely reported as issues that need to be addressed to further implement SDMX concepts and processes.

Please indicate whether your implementation(s) are directed towards[11] **Graph 2**



4. Looking Ahead: from Macro to Micro?

One of the main assets of SDMX is its information model, which describes the various concepts used in the field of official statistics. This information model is used as a foundation for standardised statistical metadata, such as the SDMX cross-domain concepts and codelists, or the SDMX Global DSDs for BoP and the SNA. These standardised statistical metadata are typically maintained in, and shared via, an

[11] Multiple answers possible.

SDMX Registry, such as the SDMX Global Registry. Hence, it is the combination of the SDMX information model, the standardised statistical metadata and its registries that make SDMX an instrumental tool for facilitating statistical data and metadata exchanges. In particular, this allows for a high degree of automation.

Yet so far the main focus of SDMX has been on aggregated, or "macro" data. However, an increasing number of members of the SDMX community in general, and of the central banking community in particular, also have to handle microdata. Indeed, the GFC underscored the importance of micro information, noting that financial stress experienced at the level of individual entities, transactions or instruments can quickly reverberate into the entire financial system. Moreover, the shortcomings of analysis based solely on macro-level data have now been recognised, and statisticians are looking into ways to better incorporate micro data to support research, economic analysis and policy making (Tissot (2016a)). Hence, a key objective for today's central bank statisticians is to take advantage of the "micro data revolution" that is currently underway.

In theory, one would like to try to use SDMX also for this purpose. But handling micro data requires some adjustments to the SDMX standards and guidelines model,[12] for example in order to exchange Register-based data in an efficient way (for instance by allowing an information model with multiple measures). There are three key aspects from this perspective.

First, one would like to be able to have, at a single entity level, multiple values possible for a given attribute, for instance when various data sources can be mobilised to identify this attribute (with the possibility of having different values). For instance, a debt security identified by its ISIN code can be characterised by several attributes such as currency of issue, location of the primary or secondary market, and governing law (Gruić and Wooldridge (2012)). The information provided by competing data providers can differ as regards the values of these specific attributes. Additional challenges relate to the incomplete information on the characteristics of many individual securities as reported by data providers.

Second, the choice of the "right" source could depend on the overall quality of this source or to other factors that could, for instance, be time-dependant: for instance, one commercial data provider would be considered as "superior" for a period of time (say, after the introduction of the euro), but not for another. Or the choice between two conflicting sources could depend on the information provided by a third data source.

A third important feature relates to aggregation. In particular, one would want to aggregate micro

[12] The issues identified below (eg mapping and transformation/consolidation issues) are not strictly SDMX-specific but rather related to the handling of micro-data; hence other standards might face similar issues too.

level information for the same units resident in a country, or for the units that have the same nationality. For instance, one would like to measure all the debt securities issued by the firms residing in a given country, and also for all the firms (resident or not) that are controlled by a parent company that resides in the specific country considered – eg to measure all the debt securities issues by Brazilian firms either through their entities located in Brazil or through their affiliates located in offshore financial centres.[13] Again, such aggregation rules could be time-dependant, for instance depending on the evolution of control relationships between individual units and their controlling parents over time.

However, any changes to existing, "core" SDMX specifications such as the information model need to be done with great care, to make sure that there are no unexpected side effects. Important work is under way by SDMX sponsor organisations such as the ECB to ensure more compatibility of SDMX with other standards used for micro-data, such as DDI (Data Documentation Initiative) and XBRL (eXtensible Business Reporting Language). In parallel, the BIS is also looking at developing a consolidated data model for the handling of its aggregated and microdata. As user and funder of Fusion Registry, the software that powers the SDMX Global Registry, the BIS can thus play an active role to contribute to the effort mentioned above. One important objective is to ensure the funding of the development of implementation tools that will assist the development and validate the outcome of the proposed changes to the SDMX information model that could, eventually, lead to a new version of the SDMX standard.

References

Bank for International Settlements (BIS) (2013): *Guidelines for reporting the BIS international banking statistics* – version incorporating Stage 1 and Stage 2 enhancements recommended by the CGFS.

——— (2015): *Introduction to BIS statistics*, BIS Quarterly Review, September.

Cook E and A D Soto (2015): *SDMX: Challenges with the adoption of evolving standards for data exchange*, United Nations Economic Commission for Europe, Conference of European Statisticians Working Paper, April.

Denk, M and W Grossmann (2011): *Towards a Best Practice of Modeling Unit of Measure and Related Statistical Metadata*, IMF Working Paper 11/152.

Gruić, B and P Wooldridge (2012): *Enhancements to the BIS debt securities statistics*, BIS Quarterly Review, December.

IFC (2016): *Central banks' use of the SDMX standard*, IFC Report, March.

International Monetary Fund (IMF) (2015): *Ninth review of the International Monetary Fund's Data Standards*

[13] For the benefit of complementing residency-based with nationality-based information, see Tissot (2016b).

Initiatives, IMF Policy Paper, May.

International Monetary Fund and Financial Stability Board (2015): *The Financial Crisis and Information Gaps – Sixth Implementation Progress Report of the G20 Data Gaps Initiative*, September.

SDMX Statistical Working group (SWG) (2016): *Content-oriented Guidelines – Introduction*, February, available on https://sdmx.org/wp-content/uploads/SDMX_COG_2016_Introduction.pdf

Tissot, B (2016a): *Closing information gaps at the global level – what micro data can bring*, in IFC Bulletin, no 41, May.

——— (2016b): *Globalisation and financial stability risks: is the residency-based approach of the national accounts old-fashioned?*, BIS Working Papers no 587, October.

Published: Volume 2018, Issue 1 / January 25, 2018