PILOT PROJECT TO IMPLEMENT SDMX-IMTS IN MOROCCO

Office des Changes (OC), Rabat, Morocco, 16-18 September 2017

[Draft Report]

Background

- SDMX, which stands for Statistical Data and Metadata eXchange, is an ISO standard designed to describe statistical data and metadata, normalize their exchange, and improve their efficient sharing across statistical and similar organizations[1]. 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.

- An inter-agency working group consisting of Eurostat, the International Trade Centre (ITC), the Organization for Economic Cooperation and Development (OECD), the United Nations Statistics Division (UNSD), and the United Nations Conference on Trade and Development (UNCTAD) was established in 2013 with the purpose of specifying uniform structures, concept definitions and code lists for the transmission of IMTS data and metadata in accordance with SDMX.

- The working group, chaired by UNSD, developed a first version of the Data Structure Definition (DSD) of SDMX standards for International Merchandise Trade Statistics (IMTS) in consultation with member countries in 2016[2].

- The working group agreed that the first version should be tested through Pilot Projects in countries with an aim to amend and improve the standards for countries' use. Morocco was one of the countries which expressed interest in the Pilot Project through formal letter exchanges between Office des Change of Morocco (OC) and UNSD in 2016. In addition, OC has created internal working group on SDMX implementation since 2015.

- To this end, UNSD and OC agreed to conduct advisory SDMX mission to Morocco from 16 to 18 October 2017 at the OC Headquarters in Rabat.

---

[1] See https://www.sdmx.org
[2] See https://comtrade.un.org/sdmx

<u>Objective</u>

a) To conduct training workshop on SDMX-IMTS to staff of OC both IT and trade statistician; the training would focus on the data structure definition version 1.0 and relevant code lists and SDMX tools currently available

b) To assist OC implementing SDMX-IMTS by conducting mapping exercise between national database and SDMX code lists

c) To make recommendation on possible amendment of current data structure definition both at technical and substantive levels.

<u>Agenda</u>

<u>Monday, 16 October 2017</u>

Morning session:

-   Welcome remarks by Head of Statistics Department of OC and project sponsor of internal SDMX working group
-   Introduction of the pilot project and adoption of work agenda (UNSD)
-   Presentation: Introduction to SDMX-IMTS – substantive and technical (UNSD)
-   Presentation: Progress made by internal SDMX working group (OC)

Afternoon session:

-   Demo: Introduction of SDMX tools and experience of mapping Morocco data and SDMX (UNSD)
-   Presentation: Overview of data compilation practices and IT platform used (OC)
-   Extraction on sample data sets used for hands-on training workshop

<u>Tuesday, 17 October 2017</u>

All day session (hands-on training):

-   Installation and setup of SDMX tools: Mapping Assistant
-   Guided hands-on training on the use of Mapping Assistant from setting up initial configuration to finalizing mapping between national code lists and SDMX code lists
-   Identifying gaps/inconsistencies between SDMX-IMTS and Morocco trade data
-   Generating data in SDMX format: the use of "Test Client" tool
-   Self-paced hands-on exercise on the use of Mapping Assistant

Morning session:

- Self-paced hands-on exercise on the use of Mapping Assistant (cont.)
- Establishing direct link between Mapping Assistant and Trade Statistics BI database

Afternoon session:

- Discussion and preparation for reports on DSD v.1.0 amendment and improvement
- Concluding meeting: summary of work done; and way forward

Picture 1: SDMX-IMTS workshop, 16-18 October 2017 at Office des Changes, Morocco

Aide Memoire

1. The internal SDMX working group at OC was established in 2015 with aims to adopt and implement international standard in statistical data exchange (SDMX) in the domains of IMTS, Balance of Payment Statistics (BOP) and Foreign Direct Investment (FDI). The

project sponsor is Mr. RHANDI Mounir and the project manager is Mr. OULJOUR Houssaine. The main tasks are as follows:

    a. Review of available documentations on SDMX

    b. Conduct analysis on data and metadata modelling in relation to existing production tools

    c. Raise awareness on SDMX to other OC staff and conduct general training on SDMX

    d. Develop and perform pilot data exchange on selected partner organizations

    e. Implement SDMX in production; and perform continuous improvement

2. As part of the tasks, OC has organised internal training on SDMX in cooperation with Haut Commission au Planning (HCP – National Statistical Office). This general training was enough to raise awareness about SDMX and to start mapping national database into SDMX DSDs; however, it was lack of training on specific domains (such as IMTS) and of the use of SDMX tools.

3. The advisory SDMX mission by UNSD aims to complement existing knowledge on SDMX by providing more detailed and thorough explanation on SDMX-IMTS Data Structure Definition and by training OC staff on the use of SDMX tools (Mapping Assistant, Test Client and Web Services) so that they can map national data and produce the data in SDMX format.

4. UNSD indicated that the pilot project aims to review current design of DSD and to test its implementability in countries. Therefore, amendment and adjustment on its structures and code lists are to be expected.

5. UNSD provided a brief introduction to SDMX-IMTS including the background, process of developing the draft DSD, its concepts, code lists and the results of the consultation related with SDMX-IMTS. SDMX was recognized and supported by the UN Statistical Commission as the preferred standard for exchange and sharing of data and metadata at its 39th session in 2008[3].

6. The current SDMX-IMTS V1.0 has 31 concepts divided into 18 dimensions (including measure dimension), 12 attributes and 1 observation. Nevertheless, based on the outcome of earlier SDMX advisory mission in Mexico; the team has used amended DSD in Morocco mission (notably the split of dimension COMMODITY_CUSTOM_BREAKDOWN into three smaller dimensions).

Notes on amendments

---

[3] See: https://unstats.un.org/unsd/statcom/39th-session/documents/statcom-2008-39th-report-E.pdf

7. OC conducted a mapping of the code lists in the DSD with codes used in Morocco trade statistics by using Mapping Assistant[4]. Below some notes and suggestion points.

TOPIC	NOTE
CL_UNIT_MEASURE	Mapping question on unit of 1000GN (gram net) which is usually used as unit of measure for gold content
CL_UNIT_MEASURE	Unit of RAUCHE (Hive) that is not available in DSD
CL_CUSTOMS_PROC	Possible divergence on reporting of WCO Revised Kyoto Convention and "Statistical Procedures"; The suggestion is to keep WCO RKC in CL_CUSTOMS_PROC; and move "Statistical Procedures" to CL_TRADE_FLOW
CL_TRADE_FLOW	To further breakdown flows into: *Export of goods after inward processing *Export of goods for outward processing *Export on intra-firm trade *Import of goods for inward processing *Import of goods after outward processing *Import on intra-firm trade
CL_TRADE_FLOW	To split goods in inward processing into with change of ownership and without change of ownership
NEW	Idea to add flag on Free Trade Agreement (FTA)/Preferential Trade Agreement (PTA) on detailed trade statistics

Next steps

8. OC and UNSD agreed on the following next steps:

OC
   a. Continue working on the implementation of SDMX-IMTS preferably by establishing and optimizing direct connection between Mapping Assistant and trade statistics database (BI)
   b. Conduct test data exchange in SDMX format with UNSD

UNSD

---

[4] See: https://sdmx.org/?page_id=4620

a. Assist and advise OC in optimizing direct data query and mapping data; and installation of SDMX web services through remote assistance tools (see tentative programme below).

Date (TBC)	Tentative Agenda
2 November 2017 10.30 EST	Optimizing direct data query; discussion on mapping; question on tools
10 November 2017 10:30 EST	Optimizing direct data query; discussion on mapping; question on tools
17 November 2017 10:30 EST	SDMX web services
3 December 2017 10:30 EST	SDMX web services
10 December 2017 10:30 EST	SDMX web services

b. Discuss with inter-agency Working Group, agree on amendments and proposed changes, and prepare updated version of DSD
c. Upgrade UNSD IT system to receive IMTS data in SDMX
d. Together with OC, prepare documentation of implementation of SDMX-IMTS in OC, and use that to develop training materials for implementation in other pilot countries

Annex I. Notes on technical assistance in implementing SDMX

The overall result of the technical assistance was a success. We could successfully assist them in generating SDMX output on their trade data. The code mapping process went through without any major problems. The tools are comprehensive in producing the output.

The SDMX tools used in the mission are available from Eurostat website and are quite comprehensive in preparing the SDMX standard trade data. Following are the tools used in assisting Morocco to produce the SDMX data: Data structure Wizard (to update DSDs), mapping assistance (to perform code mapping), test client (to query the SDMX output) and web services (to disseminate SDMX output). Out of the four tools, two tools had more significance during the mission: Mapping assistance and Test client. Morocco did not have IIS web server and hence we could not setup the web services (.net based). However, we have promised to assist them setting up the web services on apache server online.

The resulting SDMX output is attached. The work flow is detailed below:

Data preparation

The SDMX output format requires the data to be pre-prepared into certain format for straight forward mapping procedure. Normally datasets are stored in a flat model as show below. This must be transformed into the format as shown in the figure 1.

FLOW	FLOWDESC	STAT_MONTH	STAT_YEAR	CMDCODE	PARTNER1	PARTNER2	MOT	NETWEIGHT	PRIMARYVALUE	QUANTITY	UC_LIB	CMDDESC	FREQ
I	140	1	2016	101210000	GB	FR		1000	513823	2	NOMBRE	CHEVAUX REPRODUCTEURS DE RACE PURE	M
I	140	1	2016	101210000	IE	IE		500	94088	1	NOMBRE	CHEVAUX REPRODUCTEURS DE RACE PURE	M
I	140	1	2016	101299100	NL	NL		1450	35390	4	NOMBRE	CHEVAUX DE TRAIT OU DE SELLE	M
I	140	1	2016	102210000	FR	FR	1	55178	2150611	96	NOMBRE	BOVINS DOMESTIQUES, REPRODUCTEURS DE RACE PURE	M
I	140	1	2016	105111000	ES	ES	1	1864	1593809	46600	NOMBRE	COQS,POULES REPRODUCTEURS,POIDS N'EXCEDANT PAS 185G	M
I	140	1	2016	105111000	FR	FR		818	724208	21420	NOMBRE	COQS,POULES REPRODUCTEURS,POIDS N'EXCEDANT PAS 185G	M
I	140	1	2016	105120010	US	US		885	1997625	13499	NOMBRE	DINDES ET DINDONS REPRODUCTEURS,POIDS N'EXCEDANT PAS 185G	M

Figure 1 - Raw input data

Transformation

The data preparation step involves transformation of flat data (one record per series) into multiple records with number of records is equal to number of observations. This is achieved as follows. Step 1: Joins the reference tables of quantity/weight units and other required tables with the fact tables (actual data). Split the trade value into CIF or FOB based on trade flow (in case of morocco). Step 2: Add new columns called measure and obs_value that holds quantity, CIF and FOB values in rows as shown in the figure 2. However, this type of transformation is basically de-normalize the data and hence increases its size.

FLOW	FLOWDESC	MONTH	YEAR	CMDCODE	Partner1	CIF	FOB	cmdCode6Digit	cmdCodech1	cmdCodech2	cmdCodech3	measure	value	UNIT_MEASURE
I	140	11	2016	000101210000	BE	70650	NULL	HS12_010121	00	00		QTY	3	NOMBRE
I	140	11	2016	000101210000	BE	70650	NULL	HS12_010121	00	00		V_CIF	70650	MAD
I	140	11	2016	000101210000	BE	70650	NULL	HS12_010121	00	00		W_N	1500	KG
I	140	11	2016	000101299100	BE	28397	NULL	HS12_010129	91	00		QTY	3	NOMBRE
I	140	11	2016	000101299100	BE	28397	NULL	HS12_010129	91	00		V_CIF	28397	MAD
I	140	11	2016	000101299100	BE	28397	NULL	HS12_010129	91	00		W_N	1500	KG

Figure 2 - Prepared data for mapping assisstance

Mapping

Mapping assistance tool is quite flexible to perform the code mapping with the codes list of the DSD.

Output

The final output is generated through the test client. The output is also exported into CSV equivalent of the SDMX.

Size

UNSD observed that the SDMX output is approximately 9 times the size of the original data and the equivalent CSV output is close to 3 times the original data size.

Limitations

Data preparation

This step is mandatory for enabling a smooth code mapping process. The size of the dataset increased significantly due to the de-normalization procedure (as shown in figure 2). One alternative solution to reduce the data size is to normalize the data before feeding into the mapping assistance. The normalization involves two tables. First table holds all the fields except measure and obs_value. The second table (with foreign key reference to first table) holds measure and obs_values. The two tables can then be joined to produce the desired input data for mapping assistance. The join can be done in the custom query area of the mapping assistance for cleaner implementation.

Response time

The input data took a while to load due to the size of their database. Hence, UNSD proposed to maintain a new database which acts as a staging area for data preparation. The data that is to be transformed into SDMX alone is kept in this database for quicker response time.

Reusability

The mapping assistance did not allow us to r-use the previously performed mappings. We even tried to clone the mapping set but the cloned version did not allow to change the source dataset. This is recognized as a huge limitation in mapping tool as manual mapping is time consuming process (in case of morocco data). However, we could individually export each code mappings from the mapping set and imported into the new mapping sets manually.

Time period

Time period is restricted to the values available in the current datasets. For example, if the dataset is for month of December, only the 12th month value can be mapped.

Errors

Some of the errors thrown by the mapping assistance were not user friendly. For example, there was an error that says "Cannot parse column name at position 2. Please check and use an alias". We eventually figured it out that the mapping assistance will not allow input dataset column names to be "Measure", "month" or "year".

Our final proposed architecture for Morocco' SMX implementation is shown below: