PILOT PROJECT TO IMPLEMENT SDMX-IMTS IN MEXICO

INEGI, Aguascalientes, Mexico, 18-20 September 2017

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. 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.

- 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. Mexico, which was already familiar with SDMX standards and had practical experience, was one of the countries which expressed interest in the Pilot Project.

- To this end, UNSD and the National Institute of Statistics and Geography, Mexico (INEGI) agreed that the former would conduct a technical assistance mission to Mexico from 18 to 20 September 2017 at the INEGI Headquarters in Aguascalientes.

Objective

a) In the first phase, INEGI-UNSD collaboration will aim at implementing DSD version 1.0 and come up with recommendation for amendment both at technical and substantive levels.

b) Depending on the result of this phase, INEGI and UNSD may agree to fully implement the finalised IMTS-SDMX.

Opening

1. Mr Gerardo A. Durand and Mr Juan Muñoz were in charge of opening the working sessions with the objective of conceptual and technical revisions in the application of the statistical data structure definition for International Merchandise Trade Statistics of Mexico.

2. Implementation of SDMX standards in various domains is a priority for INEGI for multiple reasons. OECD recommends that its member states implement SDMX. Recently, a peer review of Statistics Production in Mexico was conducted by OECD where the importance of SDMX was stressed. Secondly, INEGI wants to transmit data to international organizations fast and with less response burden. They hope that SDMX can ensure this. Lastly, INEGI wants to make sharing of data among agencies within Mexico more efficient and faster by implementing SDMX in all domains possible.

3. INEGI had worked with OECD in the implementation of an SDMX-Dataflow for Comtrade but with a local DSD in 2011 already. INEGI also had done the comparison between SDMX-IMTS and their code lists to identify differences and gaps.

4. UNSD added that inconsistencies between SDMX standards and Mexico data were not many and it was possible to align them.

Introduction to SDMX-IMTS

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 global consultation on SDMX-IMTS.

6. 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.[1] The current SDMX-IMTS V1.0 has 31 concepts divided into 18 dimensions, 12 attributes and 1 observation. Most concepts have a standard list of possible values called code lists.

INEGI's experience of SDMX

7. Since 2004, INEGI has been working with the SDMX International Community as a member of the SDMX Technical Working Group (TWG), and the SDMX Statistical Working Group (SWG). INEGI has participated in some SDMX Pilot Projects in

---

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

collaboration with OECD and UNSD for some other domains such as Short Term Economic Statistics, NAWWE (National Accounts World Wide Exchange), Infra Annual Labour Indicators, Millennium Development Goals, and Comtrade. However, the DSD for Comtrade is a local DSD that only considers concepts for data items compiled by INEGI.

8. This experiences showed INEGI that common data exchange standards can ensure safe and efficient sharing of data, reduce response burden, and ensure accurate interpretation of data. However, the resulting data files can be large in size.

Mapping of SDMX-IMTS DSD with Mexico IMTS

9. INEGI conducted a mapping of the code lists in the DSD with codes used in Mexico trade statistics by using Mapping Assistant.2 Below discrepancies were detected:

   a. In CL_AREA, code for China and Cameroon is CM, but the internationally agreed code for China is CN, and for Cameroon CM. UNSD confirmed that this was an error which had already been corrected.
   b. In CL_UNIT_MEASURE, Barrel and Brazilian Real have the same code BRL. UNSD responded that the code for Barrel will be changed.
   c. It was noted by INEGI that CL_UNIT_MEASURE has "Pieces/items" and "Kilowatt-hours" while INEGI uses "Thousands" and "Megawatt-hours". UNSD responded that it could be addressed by UNIT_MULT.
   d. Discrepancies detected in CL_COMMODITY was found to be because of error.

10. In a follow-up discussion on the DSD between INEGI and UNSD, the following points were raised/ suggestions were made:

   a. For Economic Activity, INEGI uses NAICS (North American Industry Classification System), a standard for North American countries. They asked whether it could be added to CL_ACTIVITY. UNSD responded that the DSD had so far accommodated international standards and not regional ones. But this can be looked into.
   b. INEGI compiles Mode of Transport and Customs Procedure Code. But, after implementation of SDMX and transmission of data to UNSD, they must be consulted before UNSD publishes these information.
   c. INEGI will not use the dimension ACTIVITY. But there's no "Not Specified" category in the code list. UNSD responded that Not Specified category will be added to all code lists for concepts that are optional.
   d. If a commodity is confidential at 6, 4, 2 or altogether, it's possible to specify that by COMMODITY_1/2. So what is the function of COMMODITY_1/2_CONF?
   e. Why is the concept TIME_PERIOD_START_DATE necessary?

---

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

Proposed amendments to the DSD for SDMX-IMTS

11. INEGI and UNSD applied a trial and error approach to convert a test Mexico trade statistics dataset to conform with SDMX-IMTS. In every step/run, new impediments to implementation were detected and dealt with.

12. Below is a list of amendments made to implement SDMX-IMTS in Mexico which are also proposed amendments for the DSD. The list contains some general proposals for next version as well.

    a. MEASURE was defined as "measuredimension" instead of "dimension" in order to support the crossing of information between TIME_PERIOD and MEASURE.
    b. Instead of CL_MEASUR, MEASURE now has a "conceptScheme" named CS_MEASURE with the same codes as CL_MEASURE.
    c. In CL_UNIT_MEASURE, code for Barrel was changed to BRR. Previously, it was BRL which was a duplicate of Brazilian Real's code.
    d. _X (Not Specified) was added to CL_ACTIVITY
    e. Concept TIME_PERIOD_START_DATE was excluded because, given there are TIME_PERIOD and FREQ, this concept was deemed unnecessary.
    f. Reporting agencies should be able to choose a truncated COMMODITY_CUSTOMS_BREAKDOWN. The full 6 digits code list may be impossible for many agencies to implement. In INEGI's case, a truncated 4 digits COMMODITY_CUSTOMS_BREAKDOWN was implemented.
    g. Re-evaluate all concepts relating to Customs Breakdown (tariff line codes) to explore if less concepts and/or fewer codes can be used.
    h. Question was raised as to how UNSD can make sure they don't publish data that the sender agency doesn't want to make public. This was particularly relevant in Mexico's case because INEGI plans to grant full access to their database with specific dimensions to UNSD from where the latter will simply extract what they need and only publish the dimensions agreed to by INEGI.

13. In addition to the above, INEGI and UNSD discussed and decided on an attachment level scheme for all concepts which was not previously defined. The scheme is as below:

Id	Attachment level
FREQ	Series
TIME_PERIOD	Series
REF_AREA	Series
TRADE_FLOW	Series
COMMODITY_1	Series
COMMODITY_1_CONF	Series
COMMODITY_2	Series
COMMODITY_2_CONF	Series
COMMODITY_CUSTOM_BREAKDOWN	Series
COUNTERPART_AREA_1	Series
COUNTERPART_AREA_1_CONF	Series
COUNTERPART_AREA_2	Series

COUNTERPART_AREA_2_CONF	Series
TRANSPORT_MODE_BORDER	Series
CUSTOMS_PROC	Series
ACTIVITY	Series
TRANSFORMATION	Series
MEASURE	Observation
OBS_VALUE	Observation
UNIT_MULT	Observation
UNIT_MEASURE	Observation
COMMENT	Observation
TRADE_SYSTEM	Dataset
COMMODITY_CUSTOM_BREAKDOWN_CODE	Series
COMMODITY_CUSTOM_BREAKDOWN_DESC	Series
COUNTERPART_AREA_1_TYPE	Series
COUNTERPART_AREA_2_TYPE	Series
COUNTERPART_AREA_1_ANNOTATION	Series
COUNTERPART_AREA_2_ANNOTATION	Series
OBS_STATUS	Observation

Next steps

14. INEGI and UNSD agreed on the following next steps:

INEGI
   a. Implement the annual SDMX-IMTS flow to all data (in the work sessions a sample of the data was taken).
   b. Test performance of the SDMX Reference Infrastructure for Monthly data, and revised Monthly data flows.
   c. Assist UNSD in preparing documentation of implementation of SDMX-IMTS in INEGI.

UNSD
   a. Discuss with inter-agency Working Group, agree on amendments and proposed changes, and prepare updated version of DSD.
   b. Complete IT system to receive IMTS data in SDMX.
   c. Test performance of the system for extraction and processing of Monthly and revised Monthly data.
   d. Together with INEGI, prepare documentation of implementation of SDMX-IMTS in INEGI, and use that to develop training materials for implementation in other pilot countries.

ANNEX-I: Agenda

Monday, 18 September 2017

Morning session:

- Opening
- Introduction to the pilot project and adoption of work agenda (UNSD)
- Introduction to SDMX-IMTS – substantive and technical (UNSD)
- Experience of using SDMX– trade and other domains (INEGI)

Afternoon session:

- Discussion on SDMX-IMTS implementation in INEGI so far (challenges, lesson learnt, etc.) (INEGI)
- Identifying gaps/inconsistencies between SDMX-IMTS and Mexico trade data (UNSD)

Tuesday, 19 September 2017

All day session:

- Substantive stream: Mapping exercise between SDMX-IMTS and Mexico code lists
- Technical stream: Sharing technical expertise, exploring SDMX CSV, protocol for data exchange

Wednesday, 20 September 2017

Morning session:

- Discussion and preparation for reports on DSD v.1.0 amendment and improvement

Afternoon session:

- Concluding meeting: summary of work done; and way forward