International Seminar on the Open Data
for the Sustainable Development Goals
September 25 to 27
Seoul, Korea

SDMX and Data Interoperability

INEGI
INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA

The SDMX Initiative

Statistical Data and Metadata eXchange

INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA

Sponsors

INEGI
INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA

Features

- SDMX:
  - Is an ISO standard (ISO 17369:2013)
  - Consists of:
    1. **Technical Specifications**. Set of standards to enable the exchange of information using modern information technology
    2. **Content Oriented Guidelines (COG)**: Set of recommend practices for creating interoperable data and metadata sets using the SDMX technical standards. The guidelines focus on harmonizing specific concepts and terminology that are common to a large number of statistical domains
    3. IT architecture and software tools needed to implement the standard

Features

- Normalize the exchange of information improving the sharing across statistical organizations
  - Provides a dynamic information model to structure the data flows.
  - Describes statistical data and metadata.
  - Improve efficient sharing across statistical and similar organizations.
  - Its features are suited to support the exchange, reporting and dissemination of statistical information

Benefits

- SDMX can be used to:
  - Facilitate and define a common language for the data and metadata exchange
  - Make efficient use of technologies and standards
  - Reduce reporting burden
  - Enhance availability of statistical data and metadata for the users
  - Streamline data reporting, data dissemination and data sharing (one data is reported only once and then shared widely using modern technologies)
  - Enable the development of standard tools that can be used and shared by a wide range of national and international organizations (examples are the SDMX-RI by Eurostat, SDDS Plus by IMF, Devinfo by UNDG, etc.)

Technical Advantages of SDMX for Interoperability

- The ONS that report can automate publication of SDMX via a web server.

- Data message content can be pre-validated by SDMX data structures.

- Data can be queried (or "pulled") by collecting agencies. This avoids the creation of data packages by the reporter for each data flow (the “push” method).

- Data can be exchanged with one international organization and shared to others.

Technical Advantages of SDMX for Interoperability

- Machine to machine exchanges improves the quality and consistency of official statistics disseminated by the IOs involved.
- Shared code lists can be accessed from a registry.
- Reuse of tools for different domains.
- Feed of information from live dissemination databases.

- By using SDMX and the SDMX-RI we have perceived the following benefits:
  - We avoided to invest a significant amount of resources by having a solid system from where to start.
  - We have agreements for the support and collaboration with experienced institutions sharing our interests.

Practical experience

- There is an initial effort (conceptual and technical) to prepare the flows and to map the dissemination databases to the data structures in order to have an SDMX web service, but after that, the work is done and you will have automatized the reporting of information.

- As the exchange doesn't depend of e-mails and questionnaires that must be manually filled, you have an improvement in the opportunity and quality and a reduction of the needed work to make the report of information

Technical Working Group (SDMX-TWG)

- The SDMX-TWG is in charge of maintenance and development of the SDMX Technical Standards

  Some of the improvements that are being developed are:

  - A Validation Transformation Language (VTL) to make possible the validation of the information flows and the transformation of data
  - JSON format to make easier the process for the dissemination of information flows, making easier to present and visualize them on web browsers
  - CSV format to make easier the exchange of large amounts of data and have better integration with existent tools, and it will make easier to disseminate information too
  - Update of the Web Services RESTful API which will make easier to any application to connect directly to a SDMX Web Service an use the information

Statistical Working Group (SDMX-SWG)

- The SDMX-SWG is in charge of the development and improvement of the SDMX Content Oriented Guidelines

- For some projects both groups are collaborating to improve aspects of the standard implying technical and conceptual issues

  – An example is the work on Code-List extensions to help in the reuse of code lists, managing different versions and changes on them, and to improve the processing of very large list of codes

Conclusions

- Interoperability depends of technical and conceptual factors
- Information interoperability only can be effective if both, the source and the receptor, interpret it in the same way; in order to do that, metadata is indispensable
- SDMX has features which make it suitable to improve interoperability
  - Elements to include data and metadata to communicate the exact meaning of the facts and the quality of the process to obtain them
  - Structures to automatize data flow processing and machine to machine communication
  - Technical descriptions and formats to implement systems that can interoperate with other systems and commercial tools
  - Free support tools which can be used to support a faster and cheaper implementation

Conociendo México

01 800 111 46 34

www.inegi.org.mx
atencion.usuarios@inegi.org.mx

@inegi_informa

INEGI Informa

INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA