Pilot of loading data and metadata into SDGsdatalab

Claire Plateau
Insee- France

1 April 2021
What is the SDGs Data Lab?

- The SDGs data lab is an online platform established by UNSD
  - To allow the exchange of SDGs indicators data and metadata by countries and custodian agencies with UNSD
  - To allow comparison of national data provided by countries against data provided by international organisations
  - To gradually replace all other forms of data submission from Custodian agencies and countries to the UNSD Global SDG indicator data base

- A request of the 2017 UN Statistical Commission to present international and national data side by side in order to facilitate dialogue between countries and custodian agencies, explain discrepancies and improve coordination
Objectives of the SDGs Data Lab

- Increasing transparency of data, improving internationally comparability and data availability; make easy data validation

- Reconciling the two complementary roles of custodian agencies and countries
  
  - Custodian agencies focusing on cross-country comparability and compliance with international standards
  
  - NSOs, being the main source informing policy analysis and decision making at national level, focusing on the comparability and consistency of country-specific data published in the Global SDGs database with their own data in their national SDGs database

- But METADATA must accompany the data to explain differences that might exist. Without metadata, discrepancies will remain unexplained and misleading as there are many valid reasons why country-level data reported by international agencies and reported by countries are different.
The SDGs Data Lab: the way forward

➢ Data exchange has been established with 29 countries so far and metadata exchanges about 5 countries

➢ Active exchange with countries and stepping up data exchange with custodian agencies, although it is still only a few

➢ Up to now, only some pilot metadata exchanges

➢ Metadata Structure Definition (MSD) will be finalized at 1st 2021 trimester based on the analysis of the metadata pilot’s result

➢ In 2021, it is planned the SDG data lab to be open for expand data exchanges – data and metadata- with countries and agencies

➢ Gradually, the SDGs Data Lab will replace other forms of Custodian agencies’ data submission to the UNSD SDG platform

➢ SDMX tools such as SDMX Converter, SDMX-RI, ILO Smart, and the Metadata Authoring Tool, support SDMX data and metadata reporting to SDG Data Lab in various infrastructure environments.
Access to the datalab

- The SDG datalab is accessed by a webbrowser
  https://unstats.un.org/SDGlab
Navigation on the datalab

Navigation

1. Data Importer
2. Data Availability
3. 202 Series (Uganda, Either International or National data) Download CSV
4. Goal 1: End poverty in all its forms everywhere
   1.1.1 Proportion of population below international poverty line (*SI_POV_DAY1*)
   1.1.1 Employed population below international poverty line (*SI_POV_EMP1*)
   1.2.1 Proportion of population living below the national poverty line (*SI_POV_NAHC*)
   1.3.1 Proportion of population covered by social assistance programs (*SI_COV_SOCAST*)
   1.3.1 Proportion of population covered by social insurance programs (*SI_COV_SOCINS*)
The visual analysis page

Chart, data table, filter control by dimension, metadata to compare National data with Global ones.
During the pilot, National metadata and quite a number of UNEP and WB reference metadata sets have been posted on the SDG Data.

Further more, UNSD is actively posting global metadata to the SDG Lab using the template and authoring tool.

NY_GDP_PCAP: Annual growth rate of real GDP per capita [8.1.1]

3. Data source type and data collection method

3.a. Data sources:
- PIB: National accounts base 2014 (according to the ESA2010)
- Population: French Census survey

3.b. Data collection method:
- Real GDP is equal Current GDP serie deflated with the GDP price index base 100 in 2014; unit is Euros 2014
- Population is calculated as the average between 2 consecutive year 1st January

3.c. Data collection calendar:
- January for Population and May for GDP

3.d. Data release calendar:
- End of May; GDP data are available up to the end of the previous year (N-1 and N-2); Data for N-1 and N-2 are provisional data; Population data are provisional for the last 5 years

3.e. Data providers:
- Insee National Accounts Department and Census Department

3.f. Data compilers:
The lab is designed to receive only SDMX 2.1 structure-specific SDMX data and metadata.

- Data and metadata can be pulled from an SDMX Rest API or by dragging and dropping an SDMX file onto the page within the specific area.
France’s pilot experience: How to generate SDMX Data?

- SDGs data are stored in the French database, the BDM (Macroeconomic Data Base).
- The SDG DSD is loaded in the SDMX part of BDM.
- A CSV file is provided, mapping every serie with the SDG DSD and the codelist modality for each dimension (or attribute) of the DSD, based on their identifier (IDBANK).
- Thanks to the SDMX WebService, a SDMX-XML file can be generated, and a url to the SDMX is accessible on [https://bdm.insee.fr/series/sdmx/data/ODD_TEST](https://bdm.insee.fr/series/sdmx/data/ODD_TEST)
- The url is loaded in the SDG lab thanks to “data importer” function.
France’s pilot experience: How to generate SDMX Metadata?

- A very friendly use
- The national metadata are completed in a world template
- They are submitted to the SDMX metadata authoring tool (https://sdg-metadata-sdmx.herokuapp.com/) to convert the metadata world file into a pdf metadata for validation and a SDMX-XML Metadata
- This SDMX-XML Metadata is loaded in the SDG Datalab (data importer function)
The SDGs Data Lab: a very friendly and intuitive platform

- The SDGs Data Lab is based on SDMX dataflows. SDMX not only facilitates data exchanges but also greatly simplifies the comparison, since both the global and national datasets follow the same DSD.

- To facilitate highest rates of adoption of SDMX exchange among reporters including both countries and international agencies, the SDMX Working Group will develop an information portal with links to the DSD and MSD, technical documents, e-learning courses, tutorials, share good practices, FAQs, etc.
The information website is available at https://unstats.un.org/sdgs/iaeg-sdgs/sdmx-working-group/

SDG Data Structure Definition

The SDG Data Structure Definition can be downloaded using the links below.

- **SDG DSD**: This is the DSD in the SDMX-ML format, downloadable from the SDMX Global Registry
- **SDG DSD Matrix**: This is an Excel representation of the DSD.
- **Guidelines for use of the Global SDG DSD**: This document contains additional information about the DSD and guidelines for its usage.
- **Guidelines for the customization of the Global DSD**: This document contains information and recommendations on customization of the global SDG DSD for national use
Many thanks for your attention