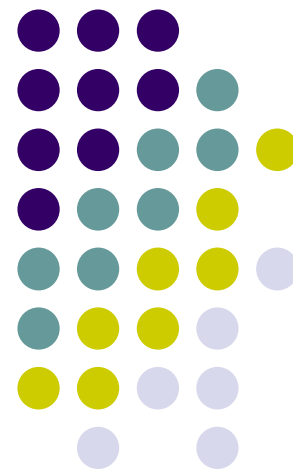


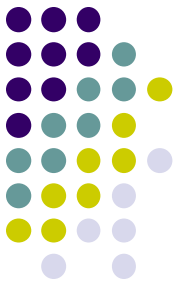
# SDMX: an Overview

Workshop on Data and Metadata  
Sharing

Bangkok, 10-14 December 2018

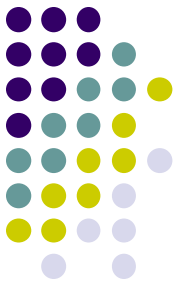
Abdulla Gozalov, UNSD





# What is SDMX?

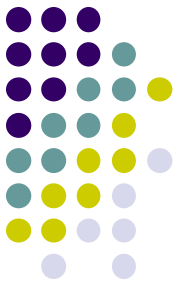
- **Statistical Data and Metadata eXchange**
- An initiative sponsored by seven international organizations
  - Bank for International Settlements
  - European Central Bank
  - Eurostat
  - International Monetary Fund
  - Organization for Economic Cooperation and Development
  - United Nations
  - World Bank



# Goals of SDMX

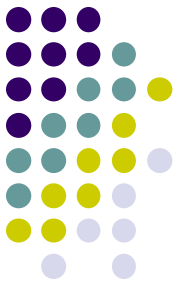
- The goal is to create common standards for exchange of statistical data and metadata to gain efficiency and avoid duplication of effort in our own work and possibly for the work of others in the field of statistical information.

# Bilateral Exchange of Statistics



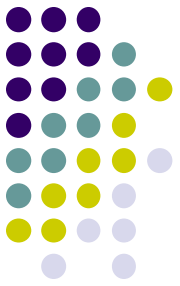
- Reporting burden on NSOs
- Collection and pre-processing burden on international institutions
- Multitude of approaches and technologies
- Every data exchange link requires a significant effort to establish.
- Established links often cannot be reused with other parties.

# Multi-lateral Data Exchange



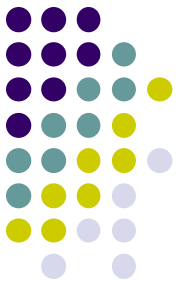
- Requires that a common standard be shared by all parties
- Only one exchange mechanism needs to be set up
- Reduces burden and duplication

# SDMX as the Infrastructure



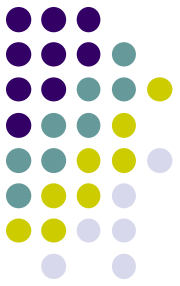
- Standards for:
  - Structuring of statistical data
  - Packaging of statistical data as XML, JSON, CSV, and other formats
  - Registry of data and metadata
  - Querying the registry
- Specifications, guidelines, tools, and manuals are freely provided to support implementation

# Statistical Data Structures



- SDMX specifies how statistical data can be structured.
  - SDMX does not detail specific structures or codes, it only provides a framework for developing those.
- **Data Structure Definitions (DSDs)** provide characteristics of the data to be exchanged.
- A DSD **must** be developed before any SDMX exchange can take place.

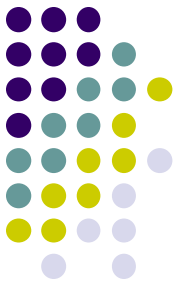
# Packaging Statistical Data



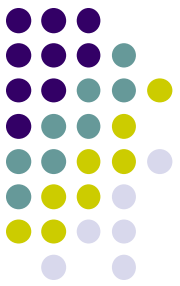
- Once a DSD has been designed, statistical data can be sent in messages structured as specified.
- SDMX provides several message formats, adapted for different scenarios.
- Validation is supported for most message formats.
- Tools are available for transformation between message types.



# Implementing SDMX Exchange



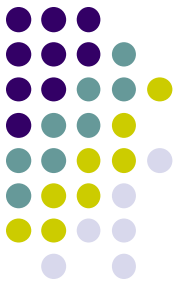
- Existing databases need not be modified.
- Software is needed (custom or off-the-shelf) to map database structures and codes to the DSD. The software retrieves data from the database and formats according to the DSD.
  - Eurostat's SDMX Reference Infrastructure allows mapping between any database and DSD
  - ILO's SMART tool can be used to map data from STATA, SPSS, and others to SDMX DSD



# Must one have a database?

- Relational databases are typically used in SDMX data exchange
  - Highly desirable to have statistical data in a database
- However, it is possible to map an Excel spreadsheet to a DSD and extract data from the spreadsheet into an SDMX file
  - It is also possible to map and convert to SDMX a flat file such as CSV
  - SDMX Converter or ILO SMART can be used

# Development of global Data Structure Definitions

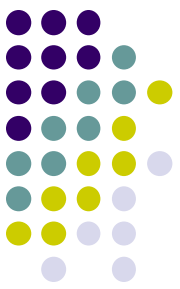


- Development of a global DSD is typically carried out by a Working Group composed of a number of international agencies and countries



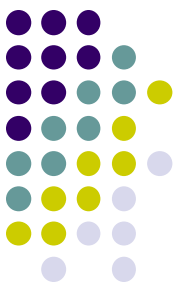
# Global DSDs

- A number of global DSDs have been developed in recent years to facilitate reporting between countries and international agencies
  - **Macro-Economic Statistics:** National Accounts, Balance of Payments, Foreign Direct Investment
    - Labor, Prices, SEEA in development or pilot stages
  - **International Merchandise Trade:** completed pilot exchange
  - **SDG Indicators:** completed pilot exchange



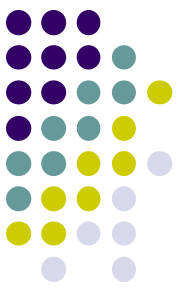
# SDMX-SDGs Working Group

- Working Group on SDMX for SDG Indicators established by IAEG-SDGs in April 2016
- First meeting in Oct 2016
- Monthly virtual meetings
- Draft DSD(s): Feb 2018
- Pilot data exchange: Apr – Sep 2018
- Official DSD and start of production data exchange: Mar 2019



# Challenges

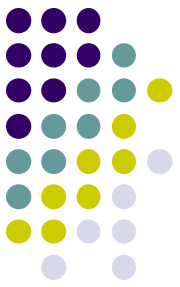
- Highly disaggregated dataset mandated by the General Assembly
  - Methodology under development for many of the indicators and breakdowns
- Guidance on specific indicator disaggregation currently under development
- SDMX solution must be customizable for dissemination by countries



# Opportunities

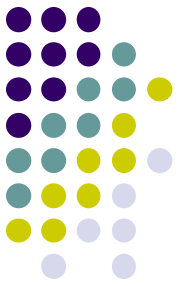
- Uniform data model
- Read and correctly interpret SDG data from any source: national, regional, global
- Uniform data access and dissemination
- SDMX-based API: JSON, XML, CSV, ...
- Greatly reduced reporting burden
- Common tools, visualizations, platforms...

# SDGs Data and Visualization Platform



- At the moment, built from data submitted by international agencies
- Pilot: exchange with international agencies as well national and regional agencies and publication of their data at the platform
  - With support provided by the UK Department for International Development





Thank you!