



**United Nations Economic Commission for Europe
Statistical Division**



Introduction to Models and Standards for Statistical Modernisation

Session 4 (ONA) : GSBPM, GSIM and SDMX

**Building national capacities on International
Standards for Official Statistics
Chisinau 19-21 September, 2022**





Contents

1. Generic Statistical Business Process Model (GSBPM)
2. Generic Statistical Information Architecture (GSIM)
3. Statistical Data and Metadata Exchange (SDMX)

modern stats

1. GSBPM

Overarching Processes							
Specify needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Reuse or build collection instruments	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Reuse or build processing and analysis components	4.2 Set up collection	5.2 Classify and code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Reuse or build dissemination components	4.3 Run collection	5.3 Review and validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production systems		5.5 Derive new variables and units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare and submit business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production systems		5.7 Calculate aggregates			
				5.8 Finalise data files			



2. What is GSBPM

Generic **S**tatistical **B**usiness **P**rocess **M**odel

- ❖ Is an international statistical standard model
- ❖ is a means to describe statistics production in a general and process-oriented way.
- ❖ is used both within and between statistical offices as a common basis for work with statistics production
- ❖ it is used for all types of data

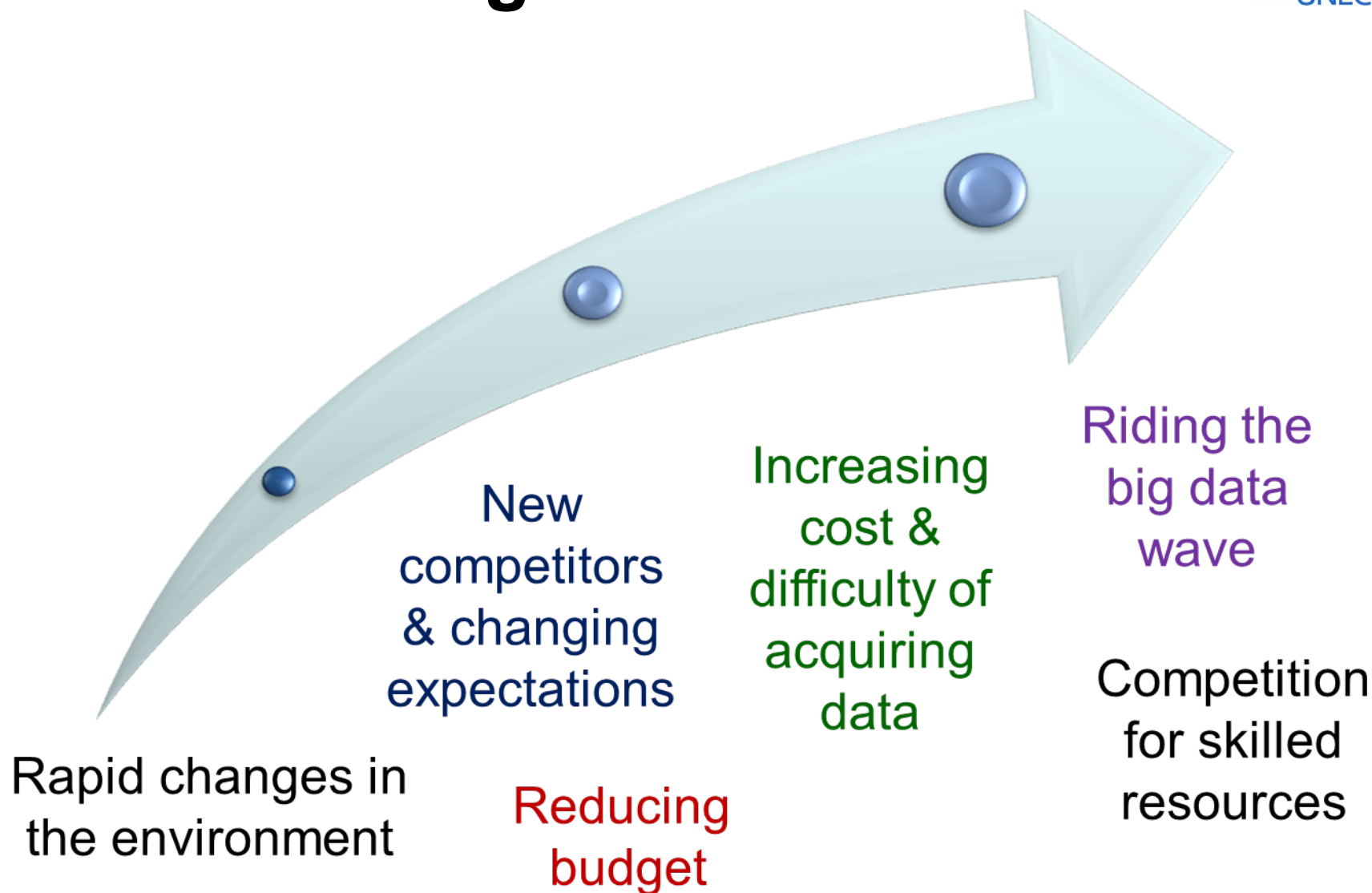


2. What is GSBPM

- ❖ Provides standard terminology to help statistical organisations:
 - Modernise statistical production processes
 - Share methods and components



The Challenges





These challenges are too big
for statistical organisations to
tackle on their own

**There is a need to work
together**



**Using common standards,
statistics can be produced
more efficiently**

No domain is special!



Why do we need the GSBPM?

- ❖ To **define and describe** statistical processes in a coherent way
- ❖ To **compare and benchmark** processes within and between organisations
- ❖ To make **better decisions** on production systems and organisation of resources
- ❖ To **support standardisation** of tools, methods and processes across statistical domains



Why do we need the GSBPM?

There are practical benefits:

- Standardisation of terminology
- Standard framework for benchmarking
- Facilitates use of common tools / methods
- Efficiency savings
- Tool for managing process quality



GSBPM applies to...

- ❖ All activities undertaken by producers of official statistics which result in data outputs
- ❖ All statistical domains
- ❖ All types of data source:
 - Surveys / censuses
 - Administrative sources / register-based statistics
 - Alternative sources (geospatial information, Big Data, ...)
- ❖ Development /maintenance of statistical registers
- ❖ National and international statistical organisations



Structure



8 Process Phases

44 Sub-processes

Overarching Processes							
Specify needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Reuse or build collection instruments	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Reuse or build processing and analysis components	4.2 Set up collection	5.2 Classify and code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Reuse or build dissemination components	4.3 Run collection	5.3 Review and validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production systems		5.5 Derive new variables and units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare and submit business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production systems		5.7 Calculate aggregates			
				5.8 Finalise data files			

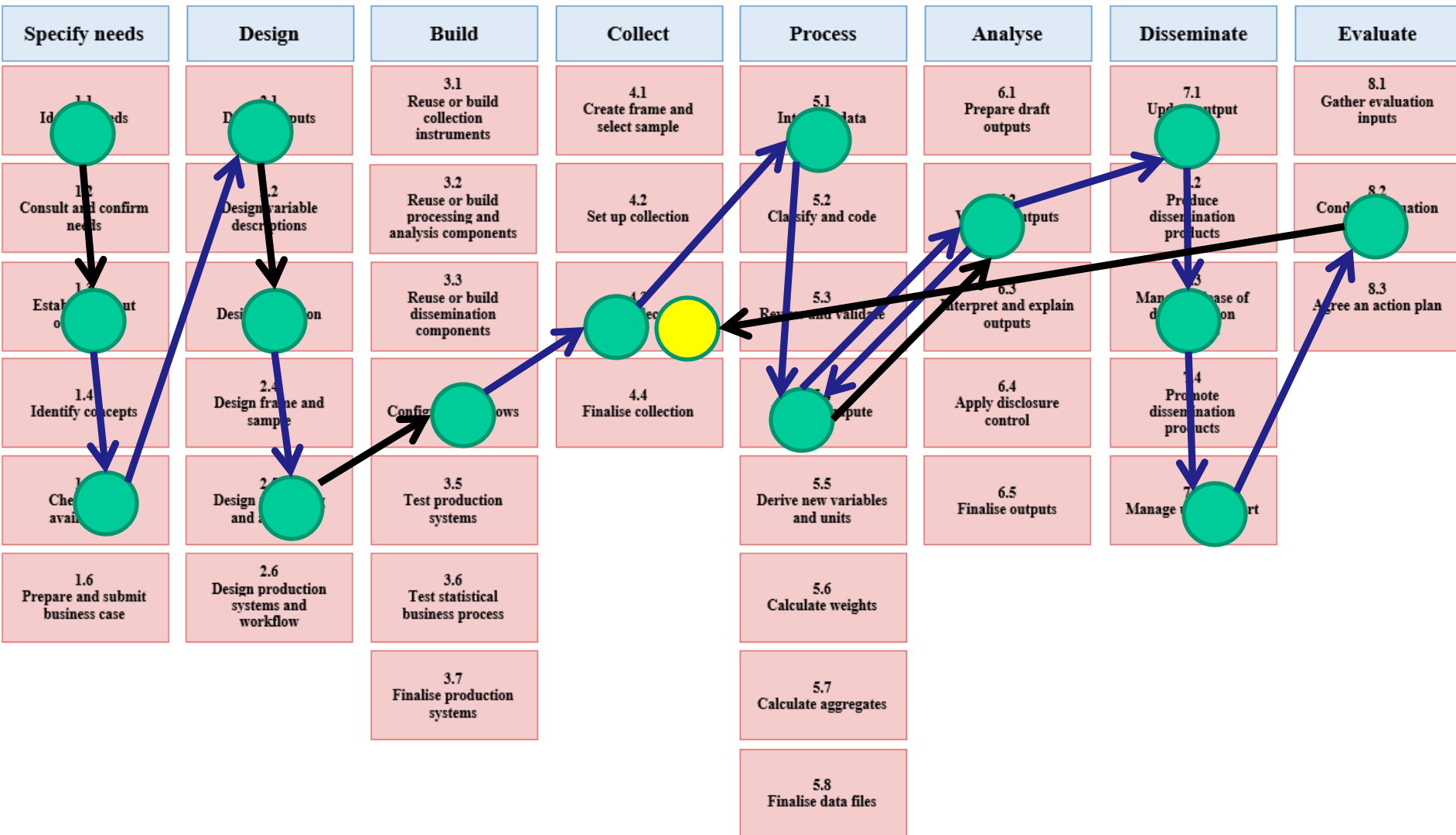
Key features

❖ **Not a linear model**

- ❖ Sub-processes are not followed in a strict order
- ❖ It is a matrix, through which there are many possible paths



Overarching Processes





**The GSBPM is used by more than
100 statistical organisations
worldwide**



Uses of the GSBPM

- ❖ Managing statistical programmes
- ❖ Cost / resource allocation
- ❖ Documenting statistical processes
- ❖ Framework for quality assessment
- ❖ Sharing statistical software

❖ **Communication!**

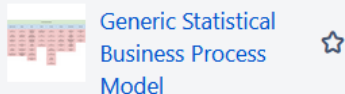
Uses of the GSBPM – in detail

- ❖ **Documentation** - “The GSBPM can provide a structure for organising and storing documentation within an organisation, promoting standardisation and the identification of good practices”
- ❖ **Process quality management** - “If a benchmarking approach to process quality assessment is to be successful, it is necessary to standardise processes as much as possible. The GSBPM provides a mechanism to facilitate this”
- ❖ **Integrating metadata and quality** - “The common framework provided by the GSBPM can help to integrate international work on statistical metadata with that on data quality by providing a common framework and common to describe the statistical business process”
- ❖ **Mapping statistical processes** - All processes that result in data outputs can be mapped to the GSBPM
- ❖ **Standardise** - terminology



UNECE

GSBPM Wiki <https://statswiki.unece.org/display/GSBPM>



Generic Statistical Business Process Model



PAGE TREE

- > GSBPM v5.1
- > Clickable GSBPM v5.1
- GSBPM v5.1 main changes
- GSBPM Resources Repository
- > Uses of GSBPM
- GSBPM Training Materials
- > Old Versions of the GSBPM
- Quality Indicators
- > GeoGSBPM
- > GSBPM Discussion Forum

Space tools



Pages 61,217 views

Edit

Save for later

Share



Generic Statistical Business Process Model

Created by Steven Vale, last modified by InKyung Choi on 08 Sep, 2021



GSBPM

Current version of GSBPM v5.1 available here



Geospatial view of GSBPM

New Geospatial-related activities and considerations needed for the production process



Clickable GSBPM

Clickable GSBPM v5.1



Learning GSBPM

Are you new to GSBPM and want to learn about it? Check out here for introductory presentations from past training and workshops



Uses of GSBPM

Use cases of GSBPM (mapping specific processes to GSBPM, using GSBPM for managing statistical programmes and etc.)



Quality Indicators for GSBPM

Quality indicators for each GSBPM subprocess



Previous versions of GSBPM

Implementations and case studies



Pages / [Generic Statistical Business Process Model](#)

GSBPM Resources Repository

Created by InKyung Choi, last modified on 08 Sep, 2021

Title	Resource type	Date	Relevant GSBPM version	Language	Author name	Author organisation	Category	Note
Mapping data production processes to GSBPM	Presentation	2014-09	5.0	English	Steven Vale	UNECE	USE CASE INTRODUCTION	GSBPM Workshop, NSO Malta, September 2014
Communicating the GSBPM – How GSBPM’s other uses can play a role	Paper			English	Joe Treacy	Central Statistics Office, Ireland	USE CASE COMMUNICATION	
Example using Statistics Norway's Business Process Model v1.1	Excel spreadsheet			English		Statistics Norway	USE CASE IMPLEMENTATION	
A five-stage data quality compliance framework	Presentation	2019-06	5.0	English	Eduardo Jallath	INEGI, Mexico	USE CASE	Presented at ModernStats World Workshop 2019 , 2019-06

Discussion forum



GSBPM Discussion Forum

Created by Steven Vale, last modified by Tetyana Kolomiyets on 06 Apr, 2021

Comments or questions about the GSBPM? Ideas for revision? Experiences to share?

Use this discussion forum to share your thoughts.

You need to be logged in to participate in the forum. To request a username and password contact support.stat@un.org



Search for Topics... 84 topics found

Show Advanced Filters

Sticky topics on top

Status	Topic	Author	# of Replies	Last Activity Date
	Issue #5: Mapping GSBPM Overarching processes to GAMS0 (from Franck Cotton) Draft mapping produced during the last GAMS0 revision	Chris Jones	0	Sep 08, 2021

More information

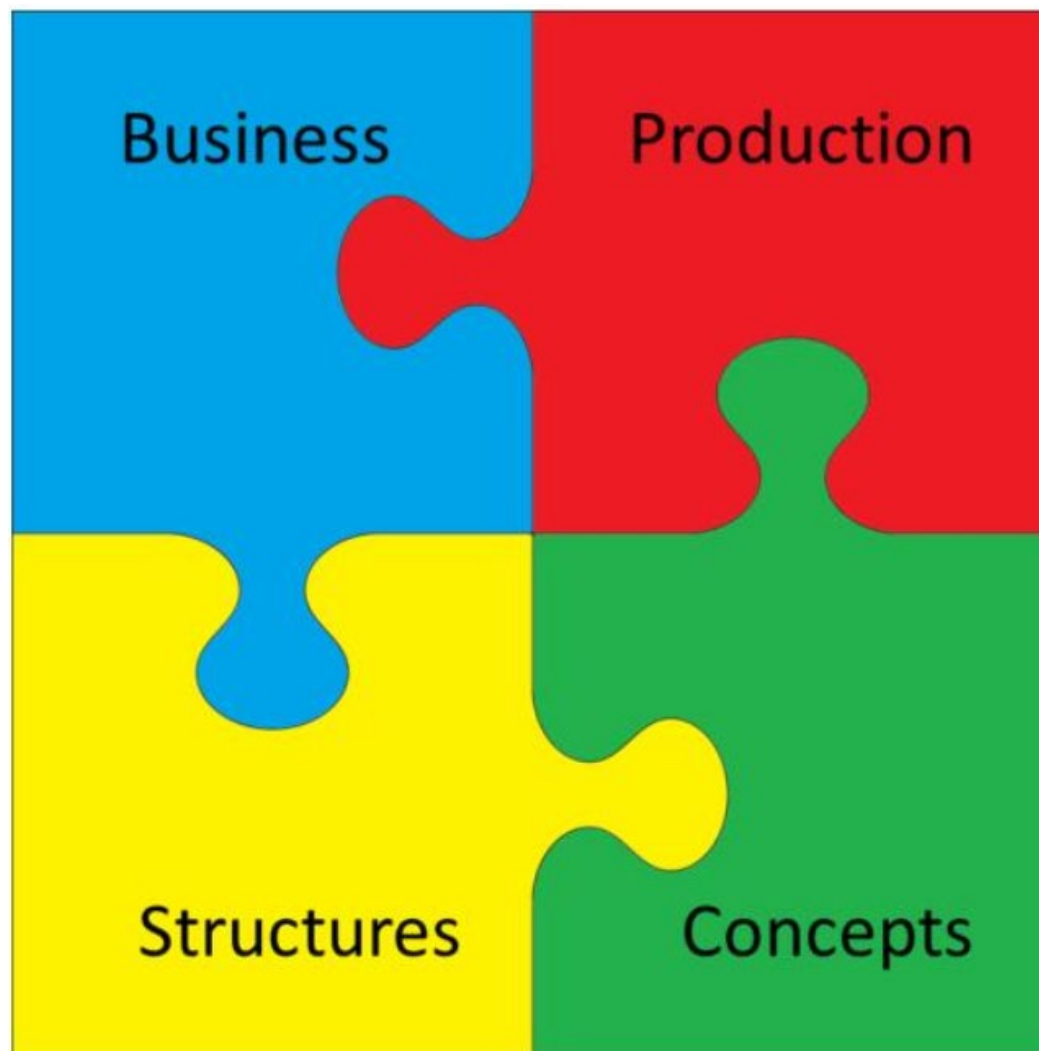


❖ GSBPM Wiki

❖ <https://statswiki.unece.org/display/GSBPM>



2. GSIM





2. GSIM

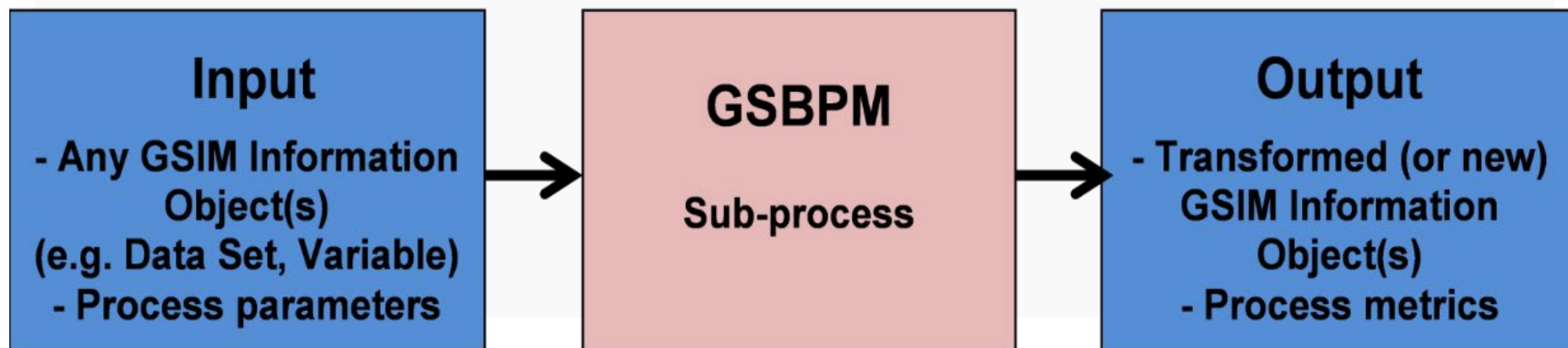
Generic Statistical Information Model

- All statistical organisations undertake similar activities but with variations in processes used.
- These activities use and produce similar information (i.e. populations for their statistical observations, statistical classifications, data sets, disseminated information).
- This information is described in different ways within each organisation.



Generic Statistical Information Model

- GSIM is the first international reference framework for statistical information
- It describes the information objects and flows within the statistical business process.





2. GSIM

- The Generic Statistical Information Model (GSIM) is the first internationally endorsed reference framework for statistical information.
- GSIM provides a set of standardised, consistently described information objects, which can be used as inputs and outputs in the design and the production of statistics.
- GSIM is a reference framework that explains relationships among the entities involved in statistical production.
- GSIM can be used to guide the development and use of consistent implementation standards or specifications.



3. GSIM

- As a common language to describe the statistical information, GSIM can facilitate communication within and between statistical organisations.
- It can provide the foundation for collaboration, standardisation and sharing of tools and methods.
- GSIM can play an important role in modernising, streamlining and, aligning standards and production associated with official statistics at both national and international levels.
- GSIM is a key element of the strategic vision of the High-Level Group for the Modernisation of Official Statistics (HLG-MOS) and is endorsed by the Conference of European Statisticians .

Examples



- ❖ Sweden
<https://statswiki.unece.org/display/GSBPM/Statistics+Sweden%3A+use+of+GSIM>
- ❖ Finland
<https://statswiki.unece.org/display/GSBPM/Statistics+Finland%3A+use+of+GSIM>
- ❖ Norway
<https://statswiki.unece.org/display/GSBPM/Statistics+Norway%3A+use+of+GSIM>

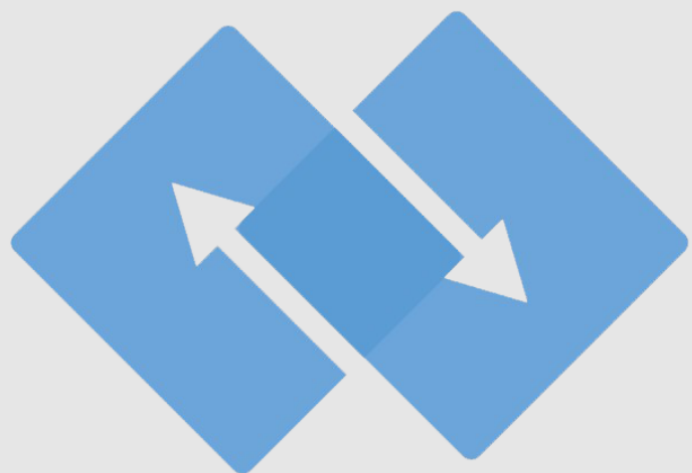
More information



❖ GSIM Wiki

❖ <https://statswiki.unece.org/display/gsim>

3. SDMX



sdmx

WHAT IS SDMX?



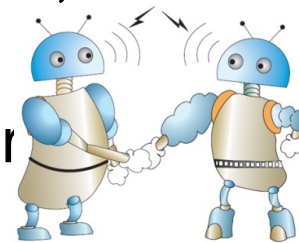
- ❖ Statistical Data and Metadata Exchange (SDMX) is the global standard for the exchange of statistical information.
- ❖ SDMX is an ISO standard adopted at level of UNSC
- ❖ SDMX is a standard for both content and technology
- ❖ SDMX facilitates data exchange between organisations and within a national statistical system
- ❖ SDMX is a common way of describing data and making it shareable
- ❖ SDMX is a set of reusable tools



WHY USE SDMX?



- ❖ To improve machine-to-machine data exchange between organisations
- ❖ To save resources – reusing data exchange systems, statistical metadata and methodology
- ❖ Improves quality – automated validations of structure and content:
- ❖ Faster and more reliable - automated workflows reduces delays from manual intervention
- ❖ Many tools freely available to use - standard software and components
- ❖ Much know-how and guidelines on how to represent and create data and metadata structures.



How does SDMX work?



- ❖ SDMX provides existing data structures for many domains to map your own data to.
- ❖ If needed you can create your own SDMX data structures using the SDMX toolkit.
- ❖ You can then use the SDMX toolkit to convert your data into SDMX

HOW DOES SDMX WORK?



The Data Structure Definition (DSD) defines:

- ❖ the **structure and valid content** of a Data Set in terms of the *Concepts* (dimensions, measures, attributes)
- ❖ **Code lists** define data concepts
- ❖ **Dimensions** - identify and describe the data.
- ❖ **Attributes** - provide additional information about the data (estimates, measures) – used in processing.

HOW DOES SDMX WORK?



❖ **Structural metadata:**

metadata used to identify, formally describe or retrieve statistical data (dimension names, variable names, dictionaries, dataset technical descriptions, dataset locations, and keywords).

❖ **Reference metadata:**

describes the contents and the statistical data quality (explanatory texts on the context of the statistical data, methodologies for data collection and data aggregation as well as quality and dissemination characteristics).

STEPS TO SHARE DATA USING SDMX



Use Case 1: Reporting National Accounts SDMX to International Organisations

You will need:

- National Accounts Global DSD (from Global Registry)
- Mapping from your database to Global DSD
- Implementation of tool such as the SDMX Converter or SDMX Reference Infrastructure (SDMX-RI) or full SDMX-based platform (.Stat suite)

How to do it:

1. Either:

- a) Fill in existing Excel questionnaire, then convert to SDMX using SDMX converter, or:
- b) Implement the SDMX-RI, create a mapping between the Global DSD and your data, or:
- c) Implement a platform, such as SIS-CC .Stat suite (or IAIS...)

If 1a) send file to IO, if 1b) or 1c) inform them of web service endpoint

STEPS TO SHARE DATA USING SDMX



Use Case 2: Collecting SDMX from another agency

1. You will need:

Agreement between exchange partners on:

- A standard reporting form (such as Global DSD)
- A means of reporting – either push (sending the file), or pull (query the data)

Map the reporting form/DSD to your database

A tool to import the SDMX into your database

2. How to do it:

- a) Agree with exchange partners on the means
- b) Create the mapping between the SDMX data and your database
- c) Use connector tools to import and remap the data ([SDMX.org/tools](https://sdmx.org/tools))

STEPS TO SHARE DATA USING SDMX



Use Case 3: Disseminating statistics

1. You will need:

- A database containing the statistical datasets
- A tool or platform that can generate SDMX from diverse data structures
- A dissemination end-point (web service)

2. How to do it:

1. Put your statistical datasets into a common database (IAIS?), or a platform such as SIS-CC .stat
2. If the common database does not have native SDMX dissemination (such as SIS-CC .Stat):
 - Implement the SDMX-RI or equivalent;
 - Connect the SDMX-RI to your database
3. Configure the SDMX web service endpoint

SDMX LINKS AND RESOURCES



- ❖ [SDMX official website](#).
- ❖ The [SDMX Content-Oriented Guidelines \(COG\)](#) recommend practices for creating interoperable data and metadata sets using the SDMX technical standards..
- ❖ [Inventory of software tools for SDMX](#) implementers and developers.
- ❖ [Eurostat main SDMX page](#).
- ❖ European Statistical System's [SDMX standards for metadata reporting](#).
- ❖ Guidelines for [managing an SDMX design project](#).
- ❖ The [SDMX Starter Kit](#) - a resource for an NSO wishing to implement SDMX.
- ❖ The [IMF SDMX Central](#) project allows users to Validate, Convert, Tabulate, and Publish data to the IMF

SDMX LINKS AND RESOURCES



- ❖ SDMX-SDG Working Group: <https://unstats.un.org/sdgs/iaeg-sdgs/sdmx-working-group/>
- ❖ A list of SDMX Tools: https://sdmx.org/?page_id=4500
- ❖ SDMX Converter: <https://circabc.europa.eu/w/browse/76a63fc2-3d22-42cc-85f2-4f6f30b4d8eb>
- ❖ SDMX RI: <https://circabc.europa.eu/w/browse/1c958330-ae5b-42e0-b7dd-3d77a0141194>
- ❖ SMART: <http://www.ilo.org/ilostat/tools/smart/index.html>
- ❖ DSD Constructor: <http://www.ilo.org/ilostat/tools/dsdConstructor/Install.htm>
- ❖ Matrix Generator: <https://gitlab.com/sis-cc/sdmx-matrix-generator>
- ❖ SDMX ISTAT Toolkit: <https://github.com/SDMXISTATTOOLKIT>
- ❖ Fusion Registry and other community tools: <https://metadatatechnology.com/community/>
- ❖ The African Development Bank [Open Data for Africa](#) project.

SDMX LINKS AND RESOURCES



Login



The Central Source of SDMX Structural Metadata

The SDMX Global Registry is the central authoritative source of SDMX structural metadata.

National or international organisations that use the SDMX standard for exchanging statistical data can register their structures with the service.

[SDMX Global Registry Home](#)

[Frequently Asked Questions](#)

The [SDMX Global Registry](#) is the technical infrastructure containing publicly available metadata material, data structure definitions (DSD), and related artefacts (concept schemes, metadata structure definitions, code lists, etc.).

SDMX LINKS AND RESOURCES



A screenshot of a web browser displaying the SDMX website. The browser's address bar shows "sdmx.org/?page_id=13". The website header includes the SDMX logo and the tagline "Statistical Data and Metadata eXchange". A navigation menu contains links for "Standards", "Guidelines", "Domains", "Implementations", "Tools", and "Events". The main content area is titled "SDMX IT tools and SDMX tutorials" and is divided into three numbered sections: "1. SDMX IT tools", "2. SDMX Tutorials", and "3. SDMX IT infrastructures". Each section contains descriptive text and links. A sidebar on the right contains a search box, a list of navigation links (Home, About us, FAQs, SDMX User Guide, Contact us), and a sign-up form for further information. The Windows taskbar at the bottom shows the Start button, system tray icons, and the time 12:07 PM.

SDMX FOR SUSTAINABLE DEVELOPMENT GOALS INDICATORS



Why use SDMX for SDG Indicators?

- ❖ A global standard Data Structure Definition, including concepts and code lists, facilitates interoperability and simplifies dissemination and reporting
- ❖ Guidance on the customization of the global DSD for national use
- ❖ A range of tools and platforms for structure maintenance, data authoring, dissemination including the Open SDG Platform, designed and optimized specifically for the dissemination of SDG indicators
- ❖ Exchange with the Global SDG Platform
- ❖ Facilitates comparisons of country data and international harmonized data as well as metadata

SDMX FOR SUSTAINABLE DEVELOPMENT GOALS INDICATORS



Open SDG - An open source, free-to-reuse platform for managing and publishing SDG data and statistics.

A screenshot of the Open SDG website. The left sidebar is dark grey with a blue header 'Open SDG' containing a search bar. The main content area is white with a blue header 'Home' and 'Edit on GitHub'. The main heading is 'Open SDG' with a colorful circular logo. Below it is the text 'Welcome to the Open SDG community' and a paragraph describing the platform. A list of key features follows, and there are links to 'Open SDG Features' and 'Frequently Asked Questions' pages.

Open SDG

Search docs

Home

About

News & events

Open SDG features

Potential users

Frequently Asked Questions

COMMUNITY

Platform updates

Support

Contribution guidelines

Code of conduct

Change log

Glossary of terms

Architecture diagram

Requirements

GitHub Next »

»Home

Edit on GitHub

Open SDG

Welcome to the Open SDG community

Open SDG is an open source, free-to-reuse platform for managing and publishing data and statistics related to the [UN Sustainable Development Goals \(SDGs\)](#). It is built exclusively with open-source libraries and tools and can be hosted and maintained using free services.

Key features include:

- Machine-readable data
- Data visualizations: graphs, data tables, and maps
- Multilingual: already available in the six official UN languages and more, using the [SDG Translations](#) resource
- Fully customisable

For more information on features see our [Open SDG Features](#) page.

Have a question? Try our [Frequently Asked Questions](#) page.

<https://open-sdg.readthedocs.io/en/latest/>


HOW TO LEARN MORE ABOUT SDMX



HOW TO LEARN MORE ABOUT SDMX



← → ↻ 🏠 https://sdmx.org/?page_id=4500 ☆ 🌐 🗄️ 🔒 🌐 🌐 🌐



The official site for the SDMX community
A global initiative to improve Statistical Data and Metadata eXchange

[SDMX Glossary](#)
[SDMX User Guide](#)
[SDMX Global Registry](#)
[SDMX Roadmap 2020](#)
[Frequently Asked Questions](#)

[Home](#) [Standards](#) [Guidelines](#) [Implementations](#) **[Tools](#)** [Learning](#) [Events](#) [Public Reviews](#) [About SDMX](#)

Software Tools for SDMX Implementers and Developers

You will find below an alphabetical list of software tools which have been developed by organisations involved in the SDMX initiative. For each tool, a succinct description is provided – you can use the respective links for a more detailed description of the tools' specific features.

Please note that Eurostat, the Statistical Office of the European Union, also maintains a webpage dedicated to [SDMX-related IT tools](#).

If your organisation is ready to offer IT tools to the SDMX user community, please contact the SDMX Secretariat (secretariat@sdmx.org) so that they can be included in the Repository.

Data Structure Wizard

Desktop application that is able to convert/edit commonly used metadata formats into SDMX-ML formats. It contains an interface that allows the user to select a given Data Structure and to complete the data according to requirements... [MORE>>](#)

Windows taskbar: Windows logo, Search, Task View, Word, File Explorer, Edge, Music, Mail, Chrome, File Explorer, PowerPoint, System tray: Network, Volume, ENG 06:11, FR 15/04/2019, Notification area: 7



Thank you for your attention....

Q&A