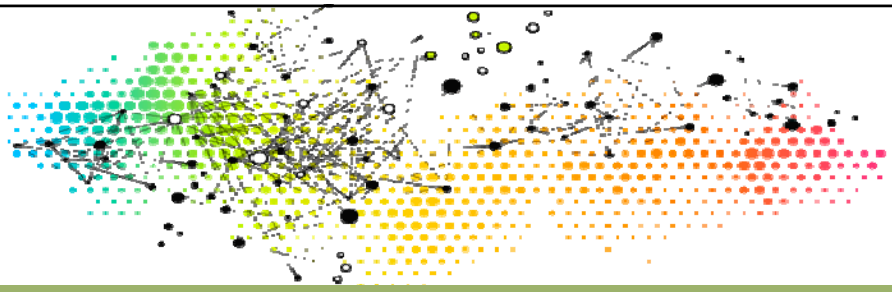


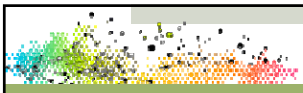
IMF Statistics Department

Generic Statistical Information Model (GSIM)

Michaela Denk, International Monetary Fund

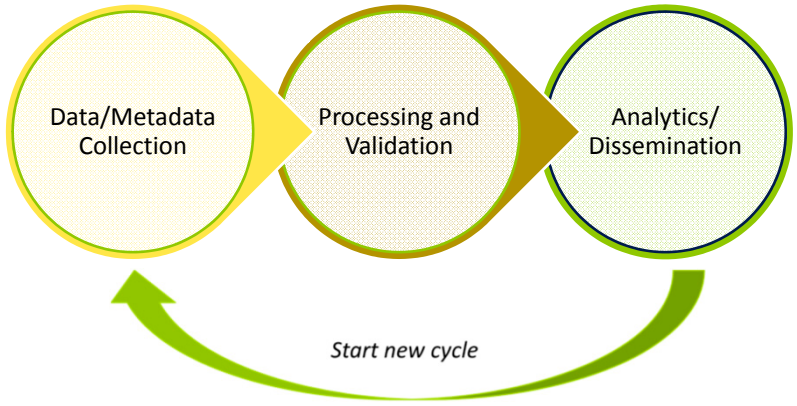
This presentation reuses material made available by the UNECE through their Virtual Standards Helpdesk at http://www1.unece.org/stat/platform/display/VSH/Virtual+Standards+Helpdesk.

The views expressed in this presentation are those of the presenter and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.



IMF Statistics Department

Statistical Data Lifecycle



- **Core process phases**
  - Every statistical organization
- **Variations**
  - Between
    - Organizations
    - Products of one organization
  - Related to
    - Terminology
    - Granularity
    - Implementation

→ **Standardized Framework: GSBPM**

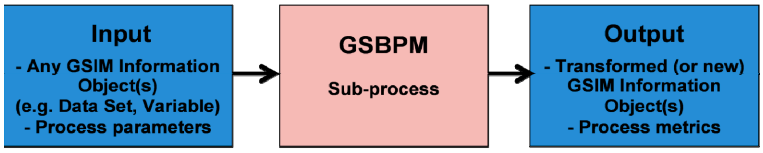
2

Generic Statistical Business Process Model (GSBPM)

- Standard framework
  - Flexible model
  - Harmonized terminology
  - Describes and defines the set of business processes needed to produce official statistics
- Helps statistical organizations
  - Move from topical stove-pipes (product-centric) to process-centric approaches
  - Define and describe statistical processes in a coherent way
  - Modernize statistical production processes
  - Compare and benchmark processes within and between organizations
  - Share methods and components

Relationship between GSIM and GSBPM

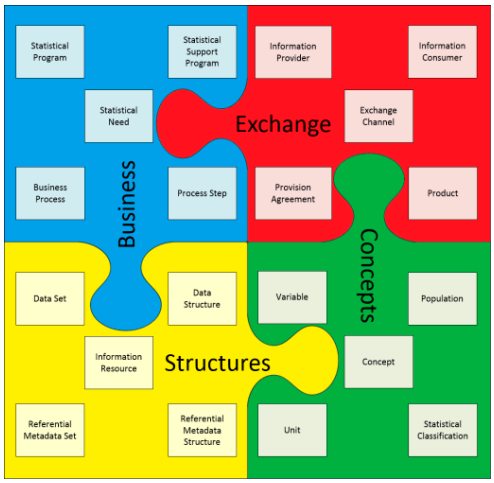
- GSBPM is about process
  - Identifies activities (sub-processes) that result in information outputs
- GSIM is about information that flows between those activities, controls activities, and documents them



- Complementary models to document the statistical business process and the information that is used in and produced by that process

Categories of GSIM Information Objects

- Business
  - Designs & plans of statistical programs, e.g. statistical program, statistical need
- Exchange
  - Incoming and outgoing information, e.g. information provider, exchange channel, product
- Structures
  - Organization and composition of data, e.g. data set, data structure, information resource
- Concepts
  - Meaning of data, e.g. variable, population, unit, statistical classification

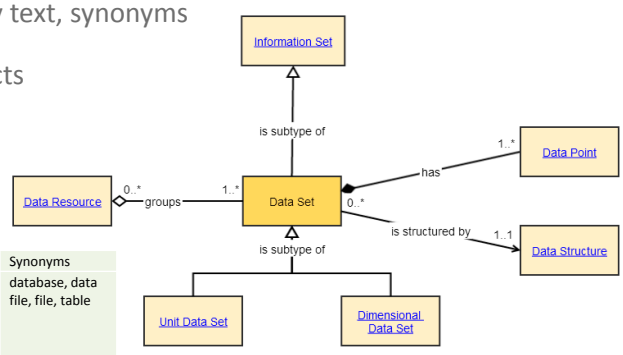


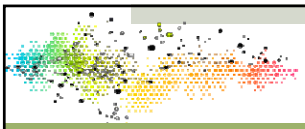
Specification of GSIM Information Objects

- Name, definition, additional explanatory text, synonyms
- Relationships to other information objects

Example: **Data Set**

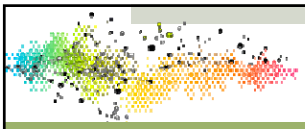
Group	Definition	Explanatory Text	Synonyms
Structures	An organized collection of data.	Examples of Data Sets could be observation registers, time series, longitudinal data, survey data, rectangular data sets, event-history data, tables, data tables, cubes, registers, hypercubes, and matrixes. A broader term for Data Set could be data. A narrower term for Data Set could be data element, data record, cell, field.	database, data file, file, table





Applicability and Use of GSIM

- GSIM is intended to apply to all information objects relevant in the production of official statistics at the national and international level
- GSIM is a reference model
  - Not all information objects are relevant to each implementation of the model
    - Differences between types of products, types of processes, organizations
- GSIM is NOT a software tool



Benefits of GSIM

- Reference framework & common terminology
  - Internationally agreed definitions, attributes, relationships describing the information used in the production of official statistics (information objects)
- Enables generic descriptions of the definition, management and use of data and metadata throughout the statistical business process → transparency
- Improves communication, coordination, cooperation, and collaboration → transparency
- Enables a higher degree of automation of the statistical business process → reproducibility
- Facilitates capacity building in statistical organizations
- Helps assess existing statistical information systems and processes

A World without GSIM

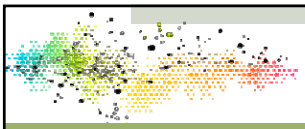
- Working in silos – each domain is different
- Lack of common terminology and understanding
- Lack of process-orientation
- Difficult to automate
- Difficult to reuse information, methods, technology
- Long time-to-market for new statistical products

One Example …

without GSIM

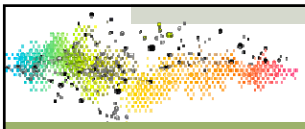
with GSIM

10



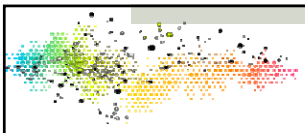
Implementing GSIM

- GSIM is a conceptual model
  - No implementation details
  - No dependence on or even reference to specific tools
- Business level
  - Map existing information model to GSIM
  - Adopt (parts of) GSIM as is
  - Combine with GSBPM
- Technical level
  - Only after business level
  - Leverage CSPA

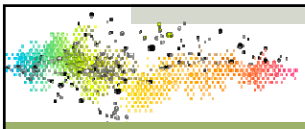


GSIM - Related Material

- "Clickable GSIM" - http://www1.unece.org/stat/platform/display/GSIMclick/Clickable+GSIM
  - Navigate the information model
  - View object definitions and relationships
- http://www1.unece.org/stat/platform/display/gsim/Generic+Statistical+Information+Model
  - Documentation of GSIM v1.1 and previous versions
  - Brochure and communication material
  - Implementation guide
  - Technical specification
  - Relation to other standards
  - Glossary
- GSBPM & GSIM case studies - http://www1.unece.org/stat/platform/display/CASES/Case+Studies+of+Metadata+use+with+GSBPM+and+GSIM



Some Background & Context

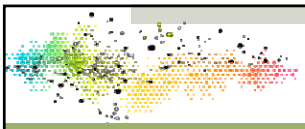


Conceptual Framework for Statistical Data Management

- Context
  - Fundamental Principals of Official Statistics (UN)
  - High-Level Group for the Modernisation of Official Statistics (UNECE HLG-MOS)
  - Modernisation Maturity Model (MMM) *
  - Data Management Maturity Assessment (e.g. Gartner)
- Generic Statistical Business Process Model (GSBPM) *
- Generic Statistical Information Model (GSIM) *
- Generic Activity Model for Statistical Organizations (GAMSO) *
- Statistical Data and Metadata Exchange (SDMX) x
- Common Statistical Production Architecture (CSPA) *
- http://www1.unece.org/stat/platform/display/VSH/Virtual+Standards+Helpdesk

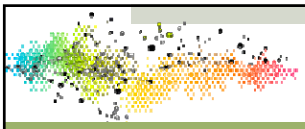
* UNECE/Conference of European Statisticians
x SDMX Initiative www.sdmx.org

14



UN Fundamental Principals of Official Statistics

- Relevance, impartiality, equal access
- Professional standards and ethics
- Accountability and ***transparency***
- Prevention of misuse
- Sources of official statistics
- Confidentiality
- Legislation
- ***National coordination***
- ***Use of international standards***
- ***International cooperation***



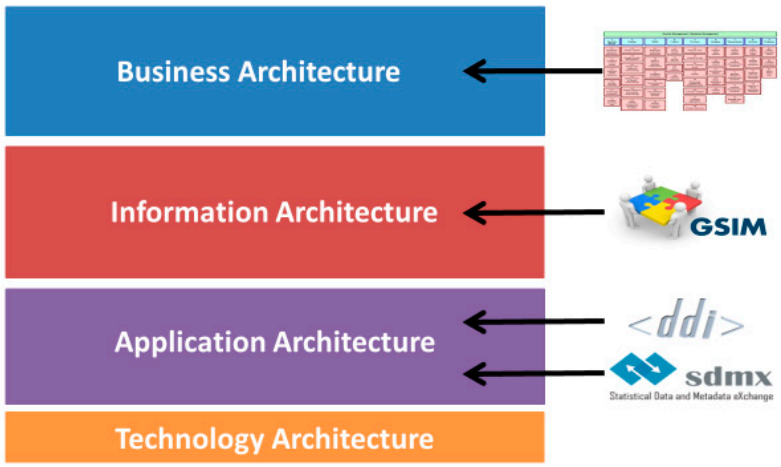
High-Level Group for the Modernisation of Official Statistics (HLG-MOS)

- Set up by UNECE Conference of European Statisticians (CES) in 2010 to oversee and coordinate international work relating to statistical modernisation
- Promotes standards-based modernization of official statistics
- Oversees modernization projects
- Manages models and tools that support modernization
- http://www1.unece.org/stat/platform/display/hlgbas/High-Level+Group+for+the+Modernisation+of+Official+Statistics

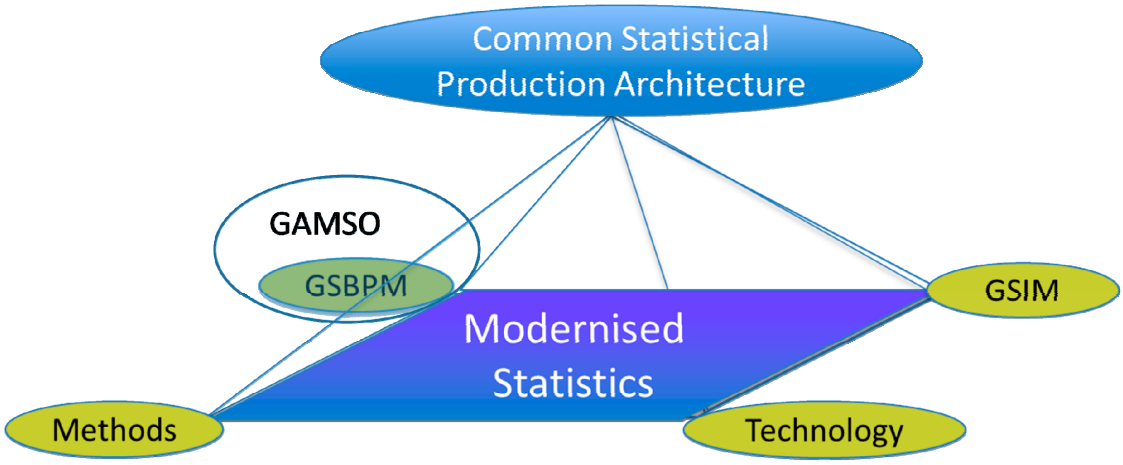
Common Statistical Production Architecture (CSPA)

- Reference architecture for statistical organizations
- Covers statistical production across processes defined by GSBPM
- Provides practical link between conceptual standards GSIM and GSBPM and technical standards such as SDMX
- Includes application architecture and technology architecture for the delivery of statistical services
- Does not prescribe technology environments
- Major aim is international collaboration & sharing
- Follows a collaborative development approach to develop re-usable systems fast and cost-effectively

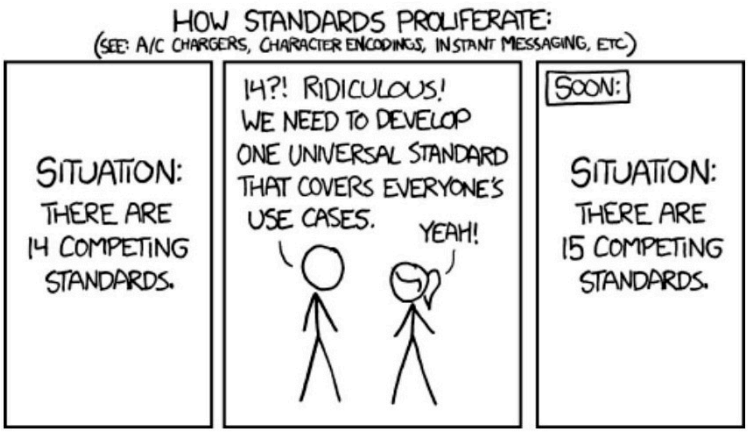
CSPA Architectural Layers



Standards in Context



How Standards Proliferate …



How Standards Proliferate …