

Generic Statistical Information Model (GSIM) - An Overview

Dan Gillman

FedCASIC – Survey Uses of Metadata
21 March 2013



Overview

- Generic Statistical Information Model
 - ▶ GSIM
 - ▶ Information or Data Model
 - What a NSO has to manage in its business
- Generic Statistical Business Process Model
 - ▶ GSBPM
 - ▶ Process Model
 - How an NSO conducts its business

Overview

- GSIM

- ▶ <http://www1.unece.org/stat/platform/display/gsim/GSIM+Development+Project>

- GSBPM

- ▶ <http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model>



What is GSIM?

- A reference framework of information objects
- It sets out definitions, attributes and relationships regarding information objects
- It aligns with relevant standards such as DDI and SDMX

GSIM Development

- UNECE Statistics Division
- Many subject matter groups
 - ▶ Information systems
 - ▶ Statistical Metadata
 - Common Metadata Framework
 - ▶ Editing
 - ▶ Disclosure
 - ▶ Dissemination
 - ▶ Many more

GSIM Development

- High Level Group
 - ▶ For Modernization of Statistical
 - Production
 - Services
 - ▶ Oversees
 - Information Systems
 - Statistical Metadata
 - GSIM
 - GSBPM

GSIM Development

■ Timeline

▶ Fall 2011

– Initial efforts with

- Australia
- Canada
- Netherlands
- Norway
- Sweden
- United Kingdom

▶ Produced simple model

▶ Realization

- Much harder problem
- More resources needed

GSIM Development

■ Timeline

▶ Winter 2012

- ABS funds project
- Project coordinator in Geneva

▶ Spring 2012

- Sprint I – 2 weeks, February, Ljubljana, Slovenia
 - Produced basic 4 part architecture
- Sprint II – 2 weeks, April, Daejeon, Korea
 - Produced basic model for each part

GSIM Development

■ Timeline

▶ Summer 2012

– Modeling Groups

- Concept Production
- Business Structure

– June – September

– Produced detailed models

▶ Fall 2012

– Integration, 1 week, The Hague, Netherlands

– Found integration points for each (6) pair

GSIM Development

■ Timeline

▶ Fall 2012, cont'd

– October –

- Comment period

– November –

- Comment resolution

– December –

- Release of Version 1.0

GSIM Development

- ▶ Winter – Spring 2013
 - Integration with other standards
 - Data Documentation initiative (DDI)
 - Statistical Data and Metadata eXchange (SDMX)
 - Neuchâtel Classification Model
 - Guided by other standards

Rationale for GSIM

- Modernisation of statistics requires:
 - ▶ reuse and sharing of
 - Methods
 - Components
 - Processes
 - Data repositories
 - ▶ definition of a shared “plug-and-play” modular component architecture
- The Generic Statistical Business Process Model (GSBPM) will help determine which components are required.
- GSIM will help to specify the interfaces.

Generic Statistical Business Process Model

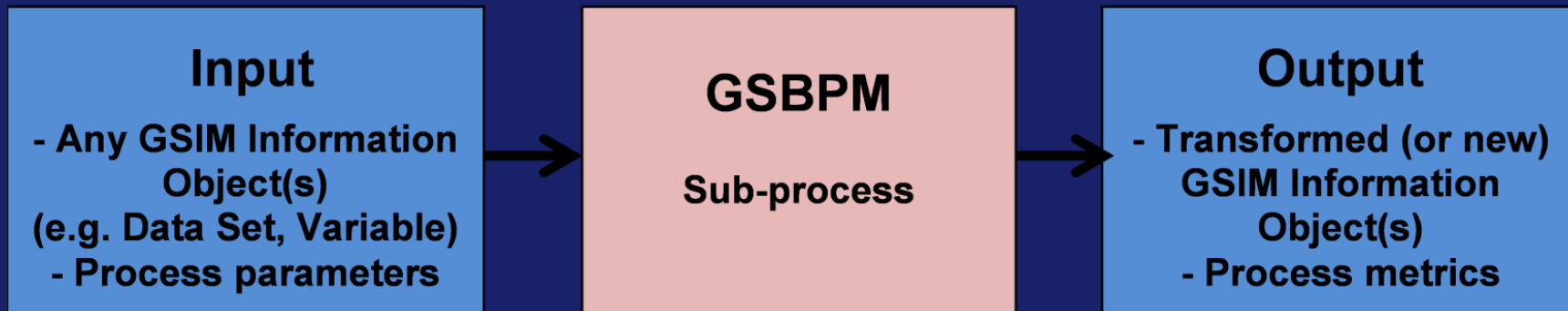
- Purpose
 - ▶ Common process model
 - ▶ Describes
 - Surveys, Censuses, Other Activities
 - ▶ Common language for survey life-cycle
 - ▶ Classification scheme for
 - Survey processes
 - Software and systems
 - Modernization efforts

Currently at BLS

- Under development
- BLS BPM
 - ▶ Based on
 - GSBPM
 - Census SLC (Survey Life-Cycle model)
 - ▶ Purpose (planned)
 - Identify system redundancy
 - Reduce costs

GSIM and GSBPM

- GSIM describes the information objects and flows within the statistical business process.



Purposes of GSIM

- Improve communication
- Generate economies of scale
- Enable greater automation
- Provide a basis for flexibility and innovation
- Build staff capability by using GSIM as a teaching aid
- Validate existing information systems



Moving to GSIM in practice

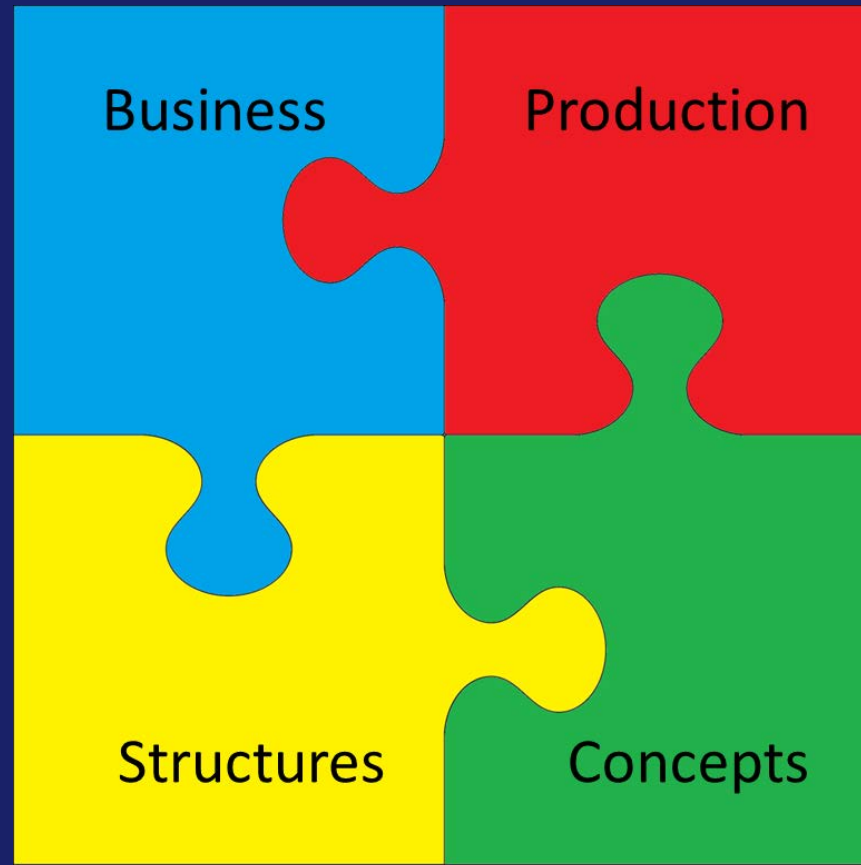
- Common terminology across and between statistical organisations.
- It allows statistical organisations and standards bodies (e.g. SDMX and DDI) to understand and map common statistical information and processes.
- The maximum benefits from standards-based modernisation will be realised if GSIM is implemented together with GSPBM.

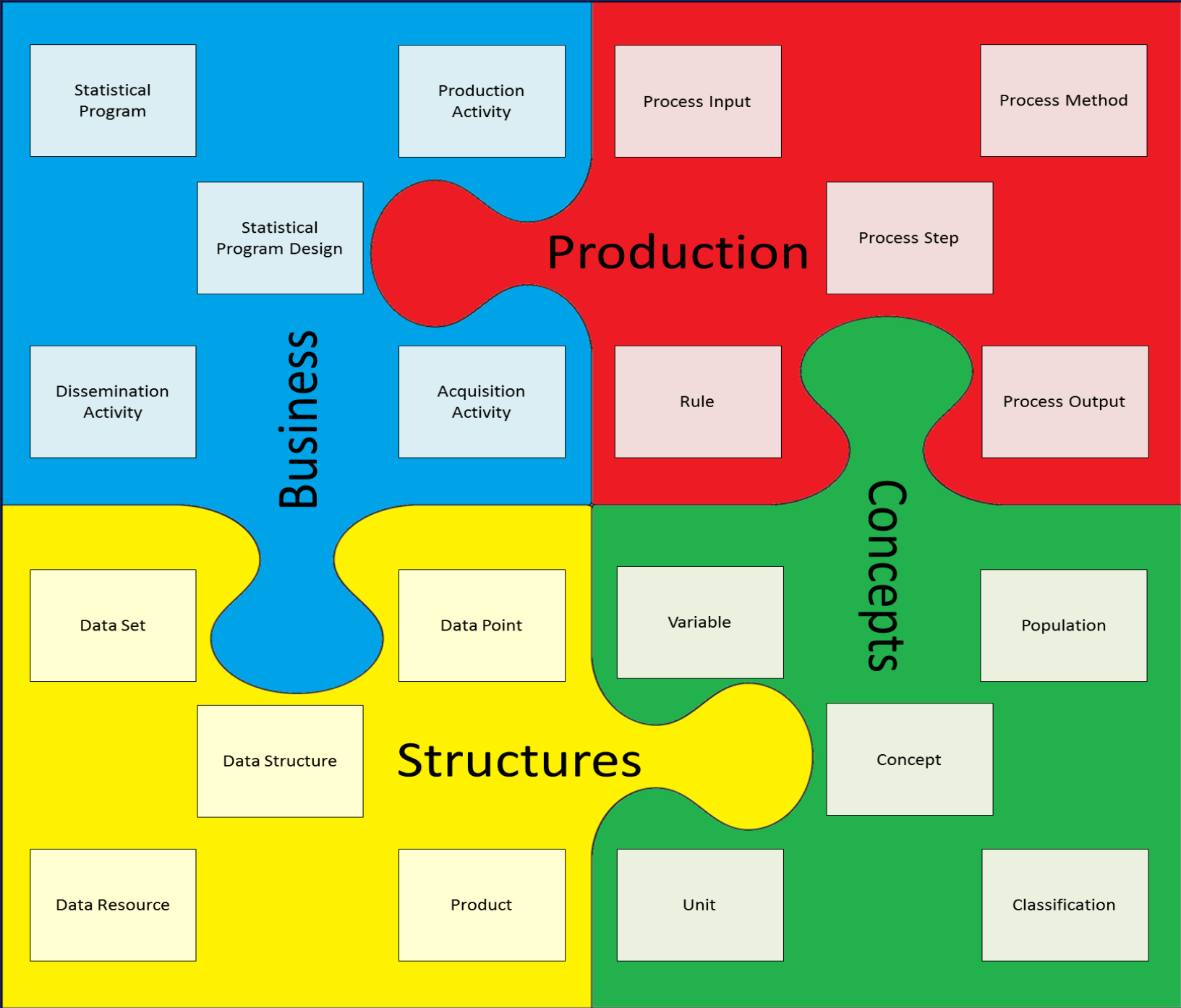
Moving to GSIM in practice

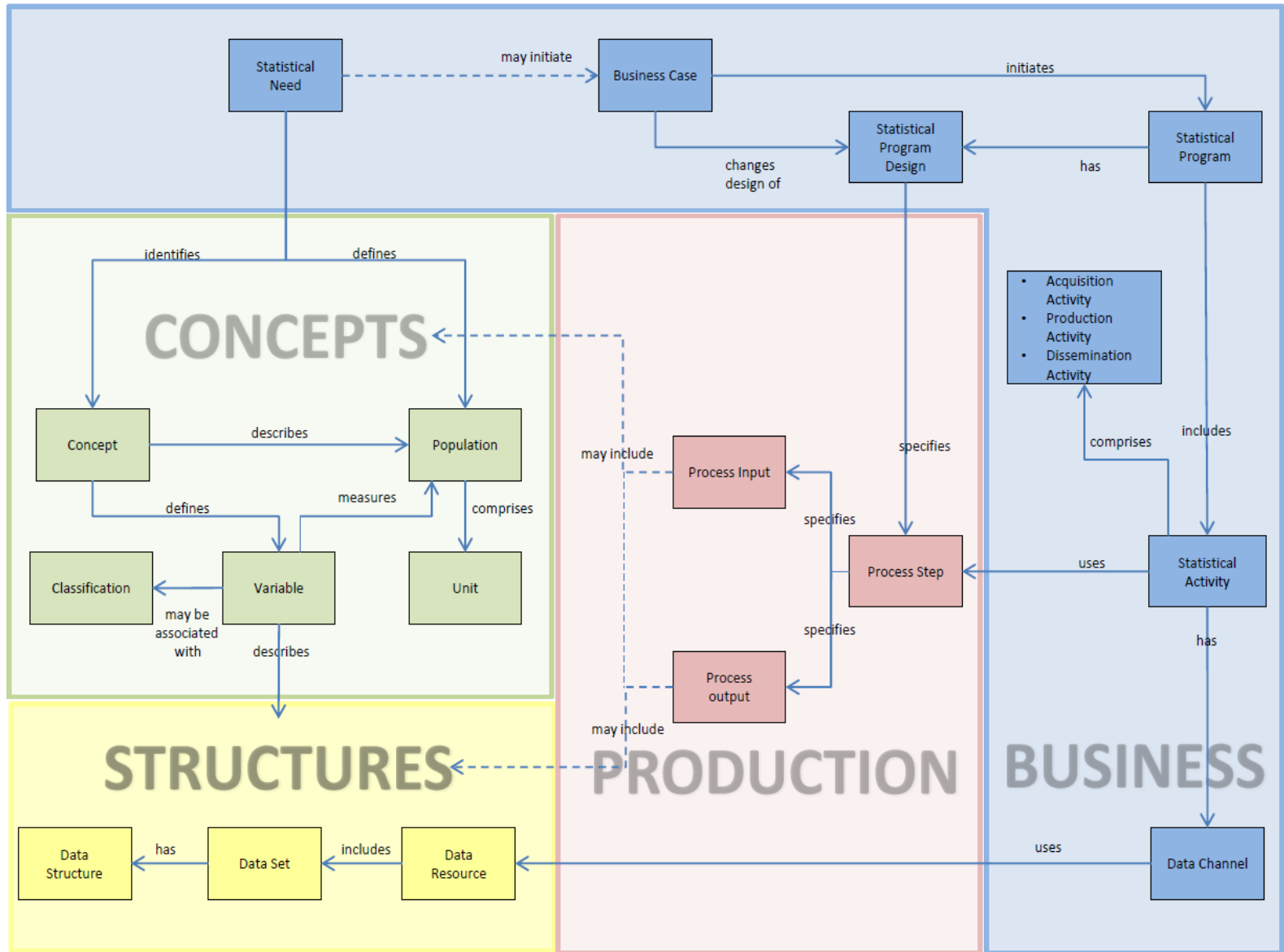
GSIM could lead to:

- A foundation for standardized statistical metadata use throughout systems
- A standardized framework for consistent and coherent design of statistical production
- Increased sharing of system components

GSIM Detail

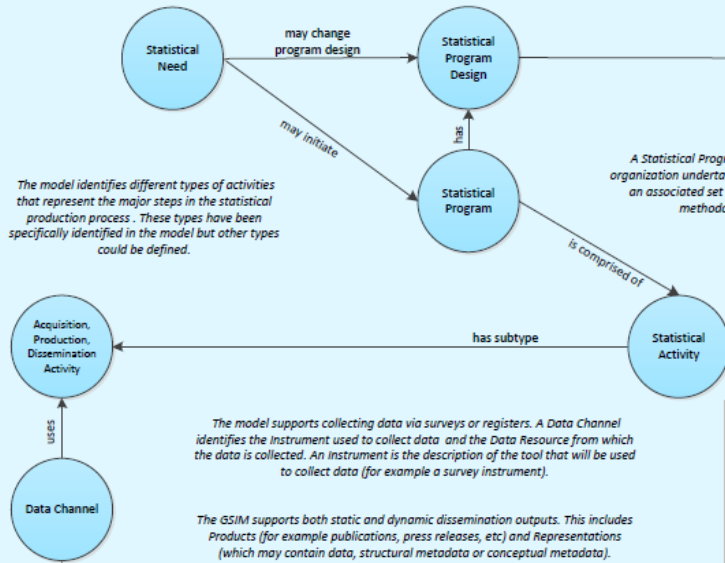




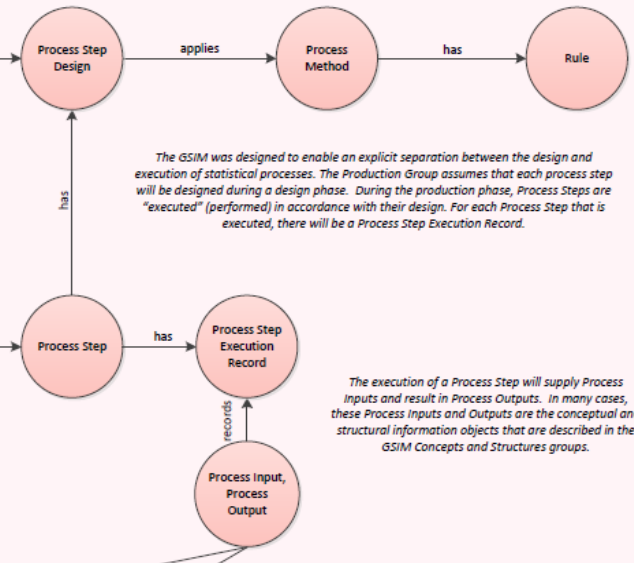


BUSINESS

The GSIM covers the whole statistical process. It includes information objects used in the Specify Needs phase (such as Statistical Need and Business Case) right through to information objects used in the Evaluation phase (such as Evaluation and Assessment).

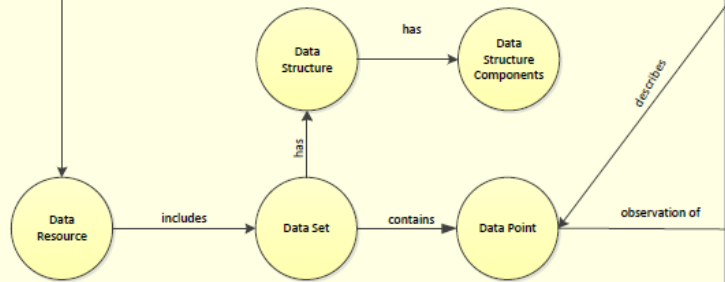


PRODUCTION



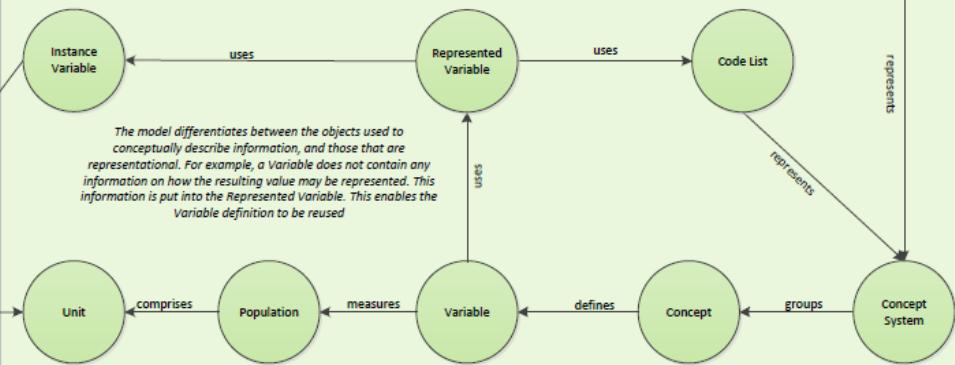
STRUCTURES

The GSIM makes a distinction between unit and dimensional data. There are a number of data structure components defined by the model. These include identifier, measure and attribute components.



CONCEPTS

The GSIM model includes rich detail on classifications. Classification management is important to statistical agencies who have complicated representations of concepts within their statistical information systems



Caveat

- Some slides taken from GSIM project web site
 - ▶ <http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=59703371>
- Slightly modified
- Some additional slides too

Contact Information

Dan Gillman

Information Scientist

Office of Survey Methods Research

www.bls.gov/osmr

202-691-7523

Gillman.Daniel@bls.gov

