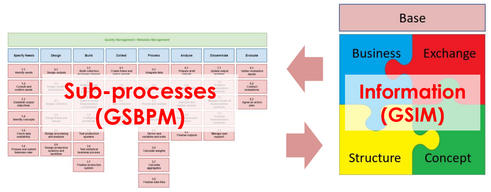
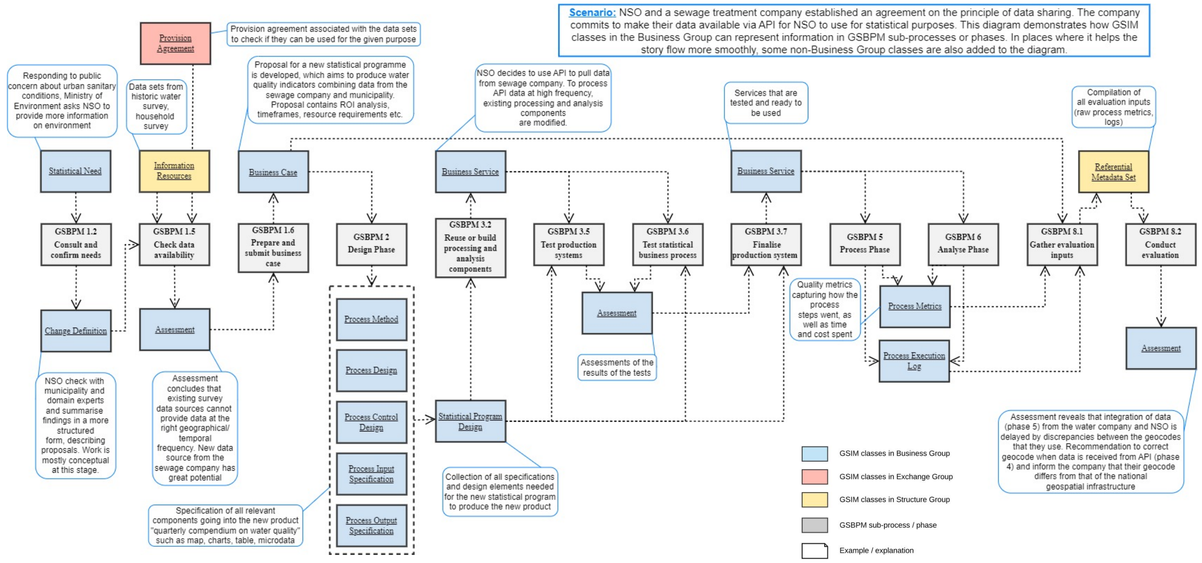
Linking metadata with business processes using modernistats by HLG - MOS models

The Generic Statistical Business Process Model (GSBPM) and the Generic Statistical Information Model (GSIM) are reference frameworks for statistical production processes and information, being used to achieve standardization and harmonization.

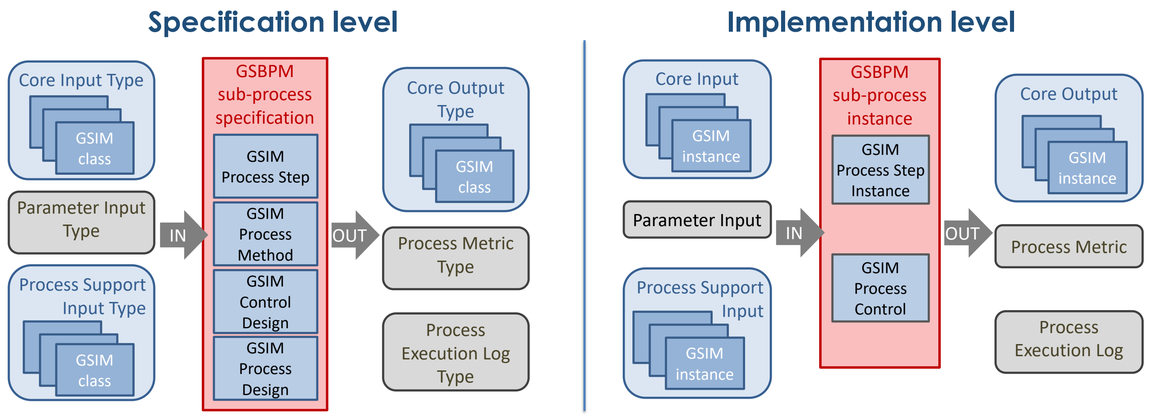
Combining these two models together helps statistical organizations to make their metadata smarter, and maximizes the benefit from their implementation for producing statistics.



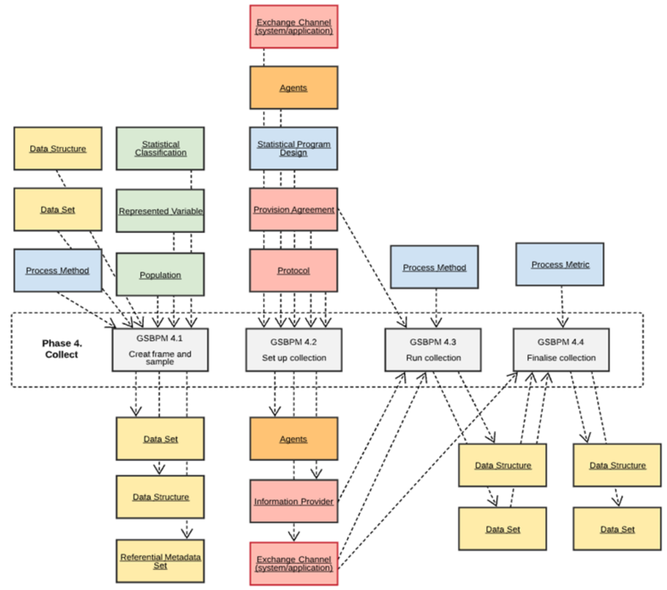
GSIM offers conceptual building blocks that systematically describe not only metadata, but also activities within the production process that GSBPM models. To illustrate this, one can consider the way that GSIM's "Business Group" relates to different GSBPM sub-processes in an example where a novel data source is being used to produce statistics.



Recent work has also used GSIM's Business Group to define the template for linking GSIM and GSBPM together, at both Specification and Implementation (Instance) levels.

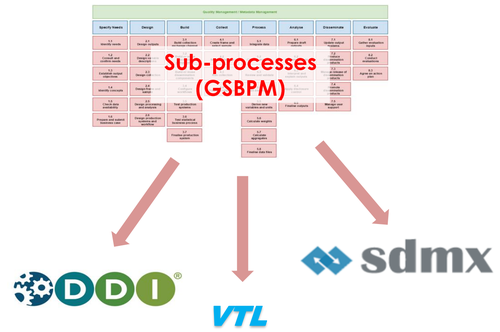


If one only considers the core input and output types at the specification level, schematic diagrams can be constructed showing the relationships between specific GSBPM sub-processes and relevant GSIM information classes. This is illustrated in the adjacent diagram, which depicts these relationships for GSBPM sub-processes within the collection (or acquisition) phase of GSBPM. Different groups of GSIM classes - business (blue), concept (green), structure (yellow), exchange (red) and base (orange) – are put together to capture the information needed to describe GSBPM sub-processes.



Linking GSBPM to implementation standards

ModernStats models can also provide common contexts and linkage points to navigate implementation standards such as SDMX and DDI. UNECE's Supporting Standards Group is currently concluding work examining different implementation standards such as these within the context of different GSBPM phases and sub-processes. An example of the output of this work is shown below for a single GSBPM sub-process.



GSBPM sub-process 1.4 (Identify concepts)

This sub-process clarifies the required concepts to be measured from the point of view of the users. At this stage, the concepts identified might not align with existing statistical standards. This alignment, and the choice or definition of the statistical and other concepts and variables to be used, takes place in sub-process 2.2 (Design variable descriptions).

SDMX Note	DDI Note
Using ConceptSchemes to organize Concepts that are to be used in a system or across an organization is a good practice that fosters data harmonization and facilitates data management activities. Relevant artifacts/instruments for this activity include Concept, ConceptScheme and Codelist; the relationship among them should be established following the principles of the SDMX Information Model. The SDMX modelling guidelines are a useful resource for this purpose, and the SDMX Glossary can also be instrumental in the data modelling process to identify the concepts and establish a common terminology and understanding.	DDI provides the means of capturing Concepts as they are identified and defined, which can then be organized (and managed) in ConceptSchemes to be used across the entire statistical process.

UNECE

Supporting standards
Modernisation of Official Statistics

Contact us at jonesc@un.org or follow us on