

A model to support the synergistic implementation of GAMSO Activities and GSBPM Overarching Processes

Developed by the HLG-MOS Supporting Standards Task Team Alignment of GSBPM Overarching Processes and GAMSO Activities

Version 1.0, 11 November 2019

About this Document

This document provides a model to support National Statistical Offices (NSOs) in having a synergistic implementation of GAMSO Activities and GSBPM Overarching Processes (OPs). The paper presents a conceptual model to identify the high-level role of the GAMSO Activities and their corresponding process activities within the GSBPM OPs. The GAMSO Activities describe generic concepts while GSBPM OPs specify activities performed across the GSBPM Phases. To illustrate the model, the paper develops a Four-Component Framework to present examples associated with the six OPs specified in GSBPM 5.1.



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>. If you re-use all or part of this work, please attribute it to the United Nations Economic Commission for Europe (UNECE), on behalf of the international statistical community.

Table of Contents

1. Introduction	3
2. Integration of GAMSO and GSBPM using a Value Chain Model	4
2.1 Value Chain Model	4
2.2 GSBPM and GAMSO	5
2.2.1 GSBPM	5
2.2.2 GAMSO	6
2.3 Integration of GSBPM and GAMSO	6
3. Current state of GSBPM 5.1 and GAMSO 1.2	7
4. A Four-Component Framework for documentation	8
5. Alignment of GAMSO and GSBPM	9
6. Implementation approaches	20
6.1 GAMSO (high-level)	21
6.2 GSBPM (low-level)	21
6.3 Synergistic	21
7. Conclusion	21
Annex A. Historical perspective of the relationship between GSBPM and GAMSO	22
Annex B. List of abbreviations	24

1. Introduction

The ModernStats models developed under the umbrella of the High-Level Group on Modernisation of Official Statistics (HLG-MOS) have had great success around the National Statistical Offices (NSOs). The Generic Statistical Business Process Model (GSBPM), released in 2008, is the first model developed to provide a common language to harmonise and standardise statistical production processes and has achieved wide adoption within the official statistical community. The Generic Activity Model for Statistical Organisations (GAMSO), released in 2017, describes and defines the four activities that take place within a typical organisation: Strategy and Leadership, Capability Development, Corporate Support and Production. GSBPM provides the details of the Production Activity of such organisation¹.

As GAMSO is gaining more attention among statistical organisations, it became of interest of statistical organisations how to use both models in an integrated way. While the 2018 revision of the two models resolved duplication and inconsistencies that used to exist, there is still no documentation on how to support synergistic adoption of GAMSO and GSBPM. This paper aims to fill the lack of documentation on how these two models can work together.

To generate this document, in 2019 the HLG-MOS set up a Task Team: *Aligning Overarching Processes in GSBPM with GAMSO under the Supporting Standards Group*. The activities mandated for this Task Team were:

“the existing works relevant to overarching processes domain that reviewed and overarching processes in GSBPM will be described in more detail, particularly those that are only mentioned (e.g. Data management, Process data management, Knowledge management, and Provider management). Then, the relationships with complementary processes activities in GAMSO would be analyzed and described. The output of the activity will be a document with explanations and examples on the overarching processes in GSBPM and GAMSO and how they relate to each other activity in GAMSO.”

We can outline the mandate as:

- Review the current description of the GSBPM Overarching Processes (OPs). Improve the narrative if needed.
- Describe those OPs depicted in the GSBPM 5.1 paper.
- Describe the interconnection of the GSBPM OPs with the GAMSO Activities. The relationship can be explicit because they have a similar name or implicit in the context of their activities.
- Include examples of the use of both.

This paper proposes a model and a framework to document GSBPM OPs and GAMSO Activities, their relationship, and examples. The framework consists of four components:

1. A description of high-level corporate activities: GAMSO Activities².
2. A description of OP activities that take place within the context of the 8 phases of GSBPM.
3. An explanation of the relationship between both.
4. An example.

The paper considers the following premises:

- Readers of this paper are familiar with the context of HLG-MOS models. They need no detailed description of GSBPM and GAMSO.
- This paper focuses on the content of GSBPM 5.1³ and GAMSO 1.2⁴ papers released in January 2019. There is no reference to the content of previous versions of GSBPM and GAMSO.

¹ Other models developed under the umbrella of the HLG-MOS can be found in <https://www.unece.org/stats/mos.html>

² This paper does not document all GAMSO activities, it only documents GAMSO Activities that have a relationship with GSBPM OPs.

³ <https://statswiki.unece.org/display/GSBPM>

⁴ <https://statswiki.unece.org/display/GAMSO>

- Readers are familiar with the naming convention of **Overarching Process** (OPs) related to GSBPM and **Activities** (upper case) related to GAMSO⁵.
- This paper develops relationships using the set of OPs in GSBPM as a point of reference. There is no relationship development for the GAMSO Activities that have no connection to GSBPM OPs. Henceforth, GAMSO activities that have a statistical component such as Management of Statistical Methodologies and Management of Consumers do not have an explicit link in this paper⁶.
- The GSBPM only has an extended description on 3 OPs. The development of specifications for other OPs was considered future work. The absence of detailed descriptions of these OPs is explained by the lack of time, not by the lack of relevance.
- The description of OPs in GSBPM includes process-level activities and high-level description to provide a corporate context where OPs are conducted. High-level descriptions are considered as “borrowed” from GAMSO.
- In this paper, the descriptions of the Activities and OPs focus on the *What*. They are generic models. The *How*, *When*, *Where* and *Who* will have to be defined in procedural manuals at each organization.

The second section of this paper provides a conceptual background on the value chain of an organisation and how this concept can conceptually link GAMSO Activities and GSBPM OPs. The third section describes a Four-Component Framework to link GAMSO and GSBPM concepts. The fourth section describes the current state of GAMSO 1.2 and GSBPM 5.1. Section five uses the Framework to document the links between GSBPM OPS and GAMSO Activities. The last section describes future work.

2. Integration of GAMSO and GSBPM using a Value Chain Model

This section provides a conceptual model to support the development of the paper. It aims to provide a clear context for the use of GAMSO Activities and GSBPM OPs and a reference framework so that users of the models can identify the scope of each model. We describe the objective of the general activities that are in GAMSO and why they are there. Likewise, we explain what level of description should be included in GSBPM OPs and how they relate to GAMSO. To provide a framework in which both models can interact, we link a value chain model to GAMSO and GSBPM.

2.1 Value Chain Model

Since the 1980s, the most common framework to conceptualise activities in an organisation has been the Value Chain Analysis developed by Michael Porter⁷. Porter’s model identifies two levels of activities: Primary and Supporting. Primary activities include all the actions that participate in the creation of a business’ offering. Support activities assist the primary activities, and they form the basis of any organisation. Figure 1 depicts Porter’s Model with Supporting Activities in horizontal bars and Primary activities in vertical bars. Supporting activities are cross-section to Primary activities.

⁵ This is just a naming convention since a process is a set of activities that interact to produce a result (Wikipedia). Therefore, GAMSO Activities can be put together into a process and Overarching Processes can be decomposed into activities.

⁶ It is expected that relationships among other Activities and OPs will be developed in future work,

⁷ Competitive Advantage: Creating and Sustaining Superior Performance (1985), Michael Porter.

Porter's Value Chain Analysis

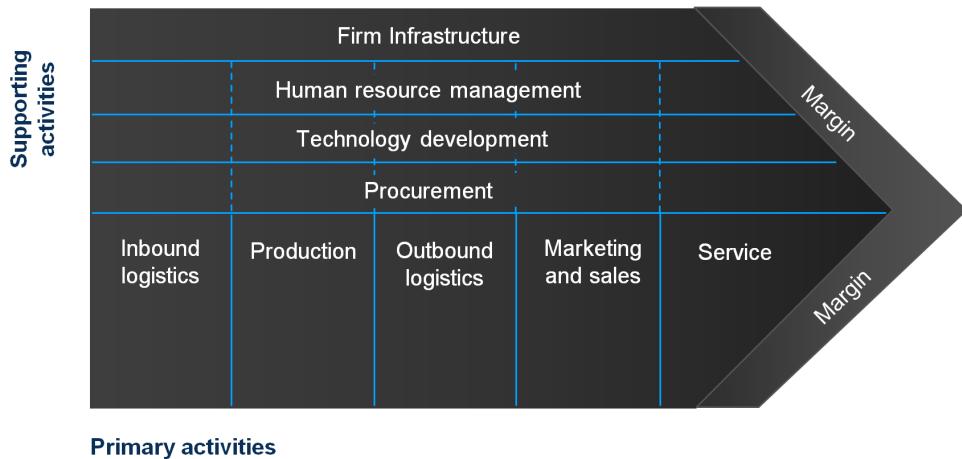


Figure 1

2.2 GSBPM and GAMSO

In 2019 the Conference of European Statisticians (CES) approved new versions of GAMSO and GSBPM. While previous versions of GSBPM and GAMSO had overlapping activities, the new versions of the models tried to make the descriptions independent. To procure independence, GAMSO 1.2 does not use the term statistical processes but Activities. This term is used to relate corporate strategies, capabilities, support, and production. In the context of Porter's model, GAMSO Support Activities are the Value Chain Support activities (horizontal bars) whereas GSBPM is in the context of the Primary activities (vertical bars).

2.2.1 GSBPM

GSBPM has 8 phases that define the statistical production: Specify needs, Design, Build, Collect, Process, Analyse, Disseminate, and Evaluate. Figure 2 illustrates these phases. Correspondingly, every phase in GSBPM includes a sequence of subprocesses.



Figure 2

The core eight phases propose a set of activities that need to be executed. They try to answer the question: What do I need to do to produce official statistics? The answer is: Specify needs, Design, Build, Collect, Process, Analyze, Disseminate, and Evaluate. It is essential to highlight that since GSBPM is a reference model, in an implementation stage, the user can adapt the phases and subprocesses to his or her specific needs. Similarly, the execution of the phases does not need to be sequential.

In addition to the eight Phases, GSBPM 5.1 includes six overarching processes:

1. Quality Management
2. Metadata Management
3. Data Management
4. Process Data Management
5. Knowledge Management

6. Provider Management

The first 3 are described in section VI of the GSBPM paper. The remaining 3 are briefly mentioned in section II, Paragraph 14. The descriptions of OPs in the GSBPM paper do not suggest any sequence of activities. They focus on highlighting a domain without providing a process context.

2.2.2 GAMSO

GAMSO is composed of four activity areas: Strategy and Leadership, Capability Development, Corporate Support, and **Production**. Each activity area is further broken down into Activities. Figure 3 shows the context of these Activities and the Production horizontal line where GSBPM is located.



Figure 3

GAMSO identifies 10 Corporate Support activities as follows:

- Manage Business Performance and Legislation.
- Manage Statistical Methodology.
- Manage Quality.
- Manage Information and Knowledge.
- Manage Consumers.
- Manage Data Suppliers.
- Manage Finances.
- Manage Human Resources.
- Manage Information Technology (IT).
- Manage Buildings and Physical Space.

Activities in GAMSO Production Activity Area are those included in GSBPM both the eight phases and the OPs⁸. Consequently, GSBPM activities are a subset of GAMSO Activities.

2.3 Integration of GSBPM and GAMSO

We can map the Value Chain, explained in section 2.1, to the production of official statistics. GAMSO refers to Support Activities, whereas GSBPM refers to Primary Activities. Nonetheless, in addition to the two traditional dimensions, Primary and Support, it is necessary to introduce a dimension to describe the OPs that support statistical production. OPs are not cross-cutting to the whole organisation; they are only cross-cutting (overarching) to the 8 phases of the GSBPM. Figure 4 shows this proposal.

⁸ From GAMSO paragraph 46: <https://statswiki.unece.org/display/GAMSO/Production>

Value Chain Model



Figure 4

Figure 5 illustrates an alternative way to depict the interrelationship between GAMSO and GSBPM using a set convention. This representation makes more evident that GSBPM Production Activities are a subset of the overall GAMSO Activities

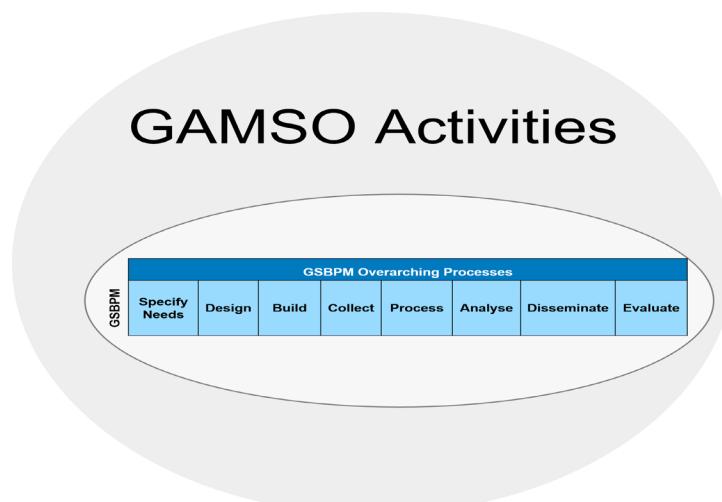


Figure 5

3. Current state of GSBPM 5.1 and GAMSO 1.2

In Section 2, we presented GAMSO as the model that should characterise the activities associated with a domain that applies to the Statistical Organisation as a whole. This implies that the definition is not restricted to statistical production. Likewise, we presented how GSBPM OPs should state the activities as they apply more specifically (although still in an over-arching manner) to the eight phases of statistical production. Regrettably, the development of the models, fully explained in Annex A, has been somehow different.

GSBPM was the first model; the authors of the model saw the need to make explicit the description of cross-section activities (or overarching) that need to be executed. However, since GAMSO did not exist, they placed these descriptions within the GSBPM paper. However, the definitions were high-level, specific activities at the Phase level were not developed.

The original intention with GAMSO was to incorporate the high-level activities described in GSBPM but the transition expected in version 1.2 did not take place. In the context of GSBPM 5.1, authors removed some OPs, but due to time restrictions, they only developed three OPs.

The task team that reviewed GSBPM 5.1 made minor changes⁹ to GAMSO to resolve inconsistencies. Despite these efforts, the relationship between Activities and OPs can fall under one of three relationship cases:

1. One to one relationship with similar identifiers, Quality Management (GSBPM) and Management of Quality (GAMSO).
2. One to one relationship with an inconsistent identifier, GSBPM 5.1 uses Provider Management, while GAMSO 1.2 uses Manage Data Supplier.
3. Many to one relationship with 4 specific terms linked to one generic term. GSBPM OPs: Metadata Management, Data Management Process, Data Management, and Knowledge Management; are linked to the GAMSO generic term Manage Information and Knowledge.

4. A Four-Component Framework for documentation

This section describes the framework to document the relationship between the GAMSO Activities and the GSBPM OPs. The intention is to depict the context under which the OPs and Activities take place in a statistical organisation. The general idea is that GAMSO Activities apply to the whole organisation while GSBPM OPs describes the activities performed at the phase/production level. Nonetheless, at every domain, GSBPM OPs need to be aligned to the specifications of the GAMSO Activity.

For instance, there can be GAMSO Activities, stated as Management of Quality policies for the entire organisation. For example, ISO 9000 can be stated as a framework for deploying quality management systems for the whole organisation. Then, specific activities should be considered at the GSBPM OP level, defining the data quality actions that take place within every phase of a statistical process. For instance, in the context of a quality management system, the OP definition should include how to collect evidence of events at phases 4 to 7. This evidence can be used later to perform the Evaluation Phase (8).

The above example illustrates the alignment with ISO 9000. The GAMSO Activities describe the general approach to quality management. OPs describe specific actions performed on every GSBPM phase.

To illustrate the operation of these two levels, we propose a Four-Component Framework with the following characteristics:

A) A high-level description of the activity - GAMSO Activity.

Support Activities (cross-cutting) required by the whole organisation should include the Supporting Activity that is already described in GAMSO or the one in GSBPM. In case no specification exists, it should be developed. The description of the GAMSO Activities not related to the GSBPM OPs (e.g. Manage Finances) is omitted.

B) Phase-level description of the GSBPM OP.

OPs descriptions should include specific activities that need to be performed at the phase level of GSBPM. It is not compulsory to include activities in every phase. In some cases, only a subset of phases might have activities.

C) Relationship.

This section should include how high-level GAMSO Activities relate to the pertinent GSBPM OP.

D) Example.

⁹ This was a minor revision for GAMSO in a second year, rather than a full revision on a fifth year.

This section should include an application of the three elements mentioned above. Thus, we need to illustrate the context of the high-level GAMSO Activity and the GSBPM OP activities at the phase level.

For every case, the objective is to depict the activities described in the corporate context and how these activities take place in the context of the 8 phases that operationalised statistical production.

5. Alignment of GAMSO and GSBPM

The alignment of GAMSO and GSBPM is mapped in this section using the model described in Section 2 and the Four-Component Framework developed in Section 3. It focuses on the description of the six domains depicted in GSBPM 5.1:

- Quality management
- Metadata management
- Data management
- Process data management
- Knowledge management
- Provider management

The cases in this section start with the high-level description. It includes text that belongs to GAMSO, and if needed, it from other sources¹⁰. Process-level descriptions of OPs, at the phase level, have not been available in any model. This paper develops these descriptions from scratch.

The domains of GAMSO Activities can be described in different schemas and sources. They can be principles (e.g. DDI Metadata), standards (e.g. ISO 9000), frameworks (e.g. NQAF) and others. To encourage consistency, every section proposes similar activities. The first four refer to GAMSO, the last two allude to GSBPM:

- Define goals and principles.
- Define strategy and plan.
- Identify requirements and business rules.
- Identify, develop, and implement a framework, including the architecture of process, data, and applications (tools).
- Deploy and manage activities.

For the sake of readability, the descriptions are not exhaustive. The intention is to provide a brief description of the domain. Those interested in getting in-depth knowledge of the subject should refer to additional sources.

Case 1: Quality Management

High-level description - GAMSO Activity - Quality Management

GSBPM includes OP Quality Management. GAMSO identifies Manage Quality. At corporate level, Quality Management activities include:

- Define Goals and Principles.
- Define Strategy and Plan.
- Identify requirements and business rules.

¹⁰ Additional references can be sourced on the GSBPM paper, other HLG-MOS models and from general sources available in the industry.

- Identify, develop, and implement a quality framework, including the architecture of process, data, and applications (tools) to support quality management.

GAMSO Quality Management Activities cover developing and administering a quality framework and cross-cutting work with tools to assure quality, e.g. compliance with the quality framework that should cover quality linked to the organisational framework, processes, and products. Such tools comprise quality indicators (including balancing quality components), user surveys, self-assessments, quality reviews or audits, certification and labelling of statistics. Quality documentation here refers to the organisational level and covers quality declarations, policies and relevant guidelines such as guidelines on the handling of errors and revisions.

References:

Section VI of GSBPM 5.1: <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1>.

<https://en.wikipedia.org/wiki/PDCA>

<https://www.bulsuk.com/2009/02/taking-first-step-with-pdca.html>

Process-level description - GSBPM OP - Quality Management

GSBPM OP needs to consider the activities defined in GAMSO and operationalise Quality Management Activities including Evaluation and Improve. The operationalisation should be at the phase level.

If the Corporate Activity defines that all processes should be under a standard or framework (e.g. ISO 9000), then the statistical process should execute actions in line with the definitions of the defined Quality Framework. These actions should be specified at the Phase level of GSBPM. Likewise, if the organisation is following the United Nations Quality Assurance Framework (UNQAF), then it should specify the actions that need to be executed at every phase to comply with the framework. These actions might include the generations of evidence that show compliance with specific quality dimensions.

If the organisation is adopting the Deming Cycle (Plan, Do, Check, and Act), then it should specify the connection with the 8 phases of GSBPM. For instance:

- Plan: Specify Needs, Design, Build.
- Do: Collect, Process, Analyze, Disseminate.
- Check: Evaluate.
- Act (Adjust): define the changes in the Plan phases.

Relationship

GAMSO defines Quality Activities that cover the whole organisation. Reference models, policies principles or standards can define the actions to be followed. GSBPM instances are a subset of the processes of the whole organisation. The OPs need to highlight the actions that need to be performed at every phase to comply with the specifications established at the GAMSO level. GSBPM operationalises GAMSO Activities at the Phase level.

Example

In a statistical organisation, GAMSO Manage Quality defines, as a Corporate Activity, a process-oriented approach, including the adoption of a Quality Management System under the guidelines of the ISO 9000 standard. The adoption implies:

- The documentation of processes and activities “do what you say and say what you do” or “how the product is fulfilling requirements”.

- The adoption of a Deming Cycle to implement continuous quality improvement.

In the context of GSBPM, there is a need to perform a mapping of the Deming cycle into the operational GSBPM phases. This mapping should include:

- The creation of specific deliverables in every phase of the GSBPM cycle (do what you say and say what you do) at expected/agreed levels of quality. For instance, there should be documentation of user needs, conceptual design, sampling design, data collection report, data analysis report, and the data set identified as the information product.
- The documentation of deviations within the Phase where they occurred so that they can be analysed in the Evaluation Phase (Check).
- The definition of the actions and timing to perform evaluations in the context of the Quality Management System defined in the ISO 9000 standard.
- The execution of business rules to obtain the relevant quality indicators on every phase to enhance the feedback for the Evaluation Phase(Check).

Case 2. Metadata Management

High-level description - GAMSO Activity - Metadata Management

Metadata Management is an OP defined in GSBPM. GAMSO includes the Activity Manage Information and Knowledge where Metadata Management is a subset. Metadata Management include:

- Define Goals and Principles.
- Define Strategy and Plan.
- Identify requirements and business rules.
- Define, develop and implement the architecture of process, data and applications (tools) to support metadata management.
- Manage Metadata Activities, including lifecycle (creation/collection, processing, storage, retrieval, archiving, and destruction).

One approach to guide the elaboration of metadata is to define a set of principles. The METIS Common Metadata Framework identifies sixteen principles:

Metadata handling

- I. Statistical Business Process Model
- II. Active not passive
- III. Reuse
- IV. Preserve Versions.

Metadata Authority

- V. Registration
- VI. Single source
- VII. One entry/update
- VIII. Standards variations

Relationship to Statistical Cycle / Processes

- IX. Integrity
- X. Matching metadata
- XI. Describe flow
- XII. Capture at source
- XIII. Exchange and use

Users

- XIV. Identify users
- XV. Different formats
- XVI. Availability

References:

A METIS description can be found in the Metadata Management description developed in GSBPM 5.1 Section VI. <https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1>.

The full text of METIS can be found in

http://www.unece.org/fileadmin/DAM/stats/publications/CMF_PartA.pdf

Wikipedia: <https://en.wikipedia.org/wiki/Metadata>

DDI: <https://ddialliance.org/>

SDMX: <https://sdmx.org/>

Process-level description - GSBPM OP - Metadata Management

Metadata are generated and processed within each phase of the model. It is requisite that metadata retains their links with data throughout the model. The following table illustrates the metadata elements required at every phase of GSBPM.

GSBPM 5.0		Metadata elements
Phases	Outputs	
1.Specify Needs	High-level objectives Relevant units Populations Universe Concepts	Conceptual metadata: Purpose, Coverage, Analysis Unit, Concepts, Universes, Variables, Classifications, Categories.
2.Design	Variables definitions Classifications Methodologies Sampling frame Collection instruments Statistical products	
3.Build	Production solutions Systems and services	
4.Collect	Collection strategy Operative reports Validation criteria Raw data sets.	Data Collection Metadata: Methodology, Sampling, Collection strategy
5.Process	Specifications for data entry, editing and coding Report on data editing Weights Anonymized micro-data	Processing metadata: Entry, Coding, Editing, Derivation, Weighting, Data discovery, Data Analysis
6.Analyse	Confidentiality statement Seasonally adjusted series Technical notes, quality indicators	
7.Disseminate	Dissemination products Metadata document	Data Distribution metadata Data Access and data availability
8.Evaluate	Evaluation report Action plan.	

Relationship

GSBPM operationalises GAMSO Activities at the phase level. GAMSO includes Metadata Management within the context of information management. These activities should include principles and standards that should be followed by the organisation. Metadata can be collected from both administrative systems and statistical systems. The latter can be related to statistical metadata standards such as DDI, Open Data and SDMX. For every standard, GSBPM needs to highlight the actions to be performed at the phase level to collect or generate specific metadata.

Example

The Data Documentation Initiative (DDI) is an international metadata standard expressed in XML and designed to describe socioeconomic surveys, censuses, and other microdata collection activities.

- Versions 1-2.5 (DDI-Codebook) - Focus on preservation of a survey. Often see the survey as a collection of data files accompanied by documentation.
- Versions 3.0-3.x (DDI-Life Cycle) - Focus on metadata reuse. Shift from the codebook centric model of early versions of DDI to a lifecycle model, providing metadata support from data study conception through analysis and repurposing of data.

These versions have different characteristics, but they identify and describe statistical studies, purpose, objectives, units, variables, code lists and questions.

GSBPM 5.0 Phases	DDI 2.0-CodeBook Metadata elements	DDI 3.2-Life Cycle Metadata elements
1. Specify Needs	Study description <ul style="list-style-type: none"> · Abstract - purposes, objectives and content · Kind of Data · Unit of Analysis · Scope · Topics Classifications · Funding 	Study Concept <ul style="list-style-type: none"> · Citation · Abstract · Purpose · Coverage topical, temporal and spatial · Analysis Unit Conceptual metadata <ul style="list-style-type: none"> · Concepts · Universes · Categories · Variables Repurposing (part) ¹¹
2. Design	Study description <ul style="list-style-type: none"> · Series Information · Universe · Coverage · Concepts · Variable definitions, universe and classifications · Source of Information · Sampling Procedure · Deviations from Sample Design 	
3. Build	<ul style="list-style-type: none"> · Questionnaires structure · Software used for data editing 	

¹¹ DDI contains the process of “Repurposing”, defined as a secondary use of a data set, or the creation of a real or virtual harmonized data set.

4. Collect	<p>Data collection</p> <ul style="list-style-type: none"> · Dates of Collection · Time Periods · Mode of Data Collection · Supervision 	<p>Data collection</p> <ul style="list-style-type: none"> · Collection strategy · Questions · Control constructs · Interviewer Instructions 	
5. Process	<p>Data Processing</p> <ul style="list-style-type: none"> · Data Editing · Other Processing 	<p>Data Processing</p> <ul style="list-style-type: none"> · Coding · Editing · Derivation · Weighting <p>Repurposing (part) ^{a/}</p>	
6. Analyse	<p>Data Appraisal</p> <ul style="list-style-type: none"> · Estimates of Sampling Error · Other Forms of Data Appraisal <p>Data Access</p> <ul style="list-style-type: none"> · Confidentiality 	<p>Data Analysis¹²</p> <ul style="list-style-type: none"> · Analysis instructions <p>Data discovery</p> <ul style="list-style-type: none"> · Variable and NCube Groups · Data Relationships 	
7. Disseminate	<p>Data Access</p> <ul style="list-style-type: none"> · Access Conditions · Citation Requirement · Disclaimer <p>Generic Dublin Core template for describing documents, programs, and other material.</p>	<p>Data distribution¹³</p> <ul style="list-style-type: none"> · Data Access description · Availability Status · Access Permission · Access type <p>Dublin Core is included as an option</p>	
Archive (Not available)	<i>Not available</i>	<p><i>Archive metadata</i></p> <ul style="list-style-type: none"> · Archival formats, locations, retention times, etc. 	
8. Evaluate	<i>Not available</i>	<i>Not available</i>	

Case 3. Data Management

High-level description - GAMSO Activity - Data Management

¹² DDI has separate phases for data discovery and data analysis, whereas these functions are combined within phase 6 (Analysis) in the GSBPM. The DDI model seems to recognize that steps such as “Data analysis” and “Repurposing” may be carried out by different organizations to the one that collected the data. This reflects a fundamental difference between practices in the research and official statistics communities.

¹³ DDI replaces the dissemination phase with “Data Distribution” which takes place before the analysis phase.

GSBPM includes OP Data Management. GAMSO includes the term Manage Information and Knowledge in this definition includes Data Management as a subset.

The term Data Management is quite broad and can include many activities related to information and data¹⁴. For the sake of simplicity, this section restricts Data Management to Data Storage Management. Data Storage Management can include:

- Define Goals and Principles.
- Define Strategy and Plan.
- Identify requirements for database technology.
- Identify, develop and implement processes, data and applications, including monitoring tools and management tools to support data storage management.
- Manage data storage activities that support the life cycle (creation/collection, processing, storage, retrieval, archiving, and destruction).

References:

Data Management Body of Knowledge (DMBOK), Data Management International www.dama.org
Storage: https://www.webopedia.com/TERM/S/storage_management.html

Process-level description - GSBPM OP - Data Management

The first two phases of GSBPM generate documents that should be managed under the content management policies. The third phase, Build, includes the development of software and probably the initial load of data. The data generated in this activity requires data storage activity to back up the software and the initial state of the data. Regular backup procedures should ensure that the system is capable of recovery in case of failure.

Operational activities in phases 4 to 8 need to be complemented with Data Storage actions to enable replicability and recovery at every phase. Likewise, there is a need to specify backup activities for the full life cycle.

Relationship

The GAMSO Data Storage Management activity should refer to the policies, principles and standards that should be followed. GSBPM at phase level should indicate when and where the data storage activities take place.

Example

In a statistical organisation, GAMSO Manage Information and Knowledge defines, as a Corporate Activity, includes the description of production data including:

- The documentation of processes and activities “do what you say and say what you do” or “how the product is fulfilling requirements”.
- The adoption of a Deming Cycle to implement continuous quality improvement.

In the context of GSBPM, there is a need to perform a mapping of the Deming cycle into the operational GSBPM phases. This mapping should include:

¹⁴ For instance, Data Management Body of Knowledge (DMBOK) includes the following knowledge areas: Data Handling, Data Ethics, Data Governance, Data Architecture, Data Modeling and Design, Data Storage and Operations, Data Security, Data Integration and Operability, Document and Content Management, Reference and Master Data, Data Warehousing and Business Intelligence, Metadata Management, Data Quality, Big Data and Data Science, Data Management, and Maturity Assessment.

- The creation of specific deliverables in every phase of the GSBPM cycle (do what you say and say what you do) at expected/agreed levels of quality. For instance, there should be documentation of user needs, conceptual design, sampling design, data collection report, data analysis report, and the data set identified as the information product.

Case 4. Process Data Management

High-level description - GAMSO Activity - Process Data Management

GSBPM identifies Process Data Management as an OP. GAMSO includes the term Manage Information and Knowledge. In the description, it includes Data Management that implicitly includes Process Data Management as a subset. Process Data Management include:

- Define goals and principles.
- Define Strategy and Plan.
- Identify requirements and business rules.
- Define, develop and implement the architecture of process, data and applications (tools) to collect process data.
- Manage Process data activities, including lifecycle (creation/collection, processing, storage, retrieval, archiving, and destruction).

Process Data Management includes activities of registering, systematising and using data about the implementation of the statistical business process. Process data can aid in detecting and understanding patterns in the data collected, as well as in evaluating the execution of the statistical business process.

Process Data Management is sometimes identified as Paradata management.

References:

<https://en.wikipedia.org/wiki/Paradata>

Process-level description - GSBPM OP - Process Data Management

Every process creates data during the execution of the program cycles. The first 3 phases do not produce this type of information. Phases 4 to 7 have computer-based execution and therefore are subject to generate data that tracks the execution.

Relationship

GAMSO defines Process Data Management Activities cover the whole organisation. GSBPM instances are a subset of the processes of the whole organisation. The OPs need to highlight the actions that need to be performed at every phase to comply with the specifications established at the GAMSO level. GSBPM should make explicit the paradata that needs to be generated at the GSBPM phase level.

Example

The GAMSO level requires a general policy that defines the characteristics and requirements of the processes of every organisation. For every process, there is a need to define provider, input, process, controls, output, evidence, users, roles and actors. The core of this process is evidence and paradata.

Case 5. Knowledge Management

High-level description - GAMSO Activity - Knowledge Management

GSBPM includes the term Knowledge Management. GAMSO uses Manage Information and Knowledge. Knowledge Management is a subset of Manage Information and knowledge. Knowledge management usually refers to the management of content. This section operationalises Knowledge Management as Document and Content Management. It includes:

- Define Goals and Principles.
- Define Strategy and Plan.
- Identify requirements, legislation and business rules.
- Define, develop and implement the architecture of process, data and applications (tools) to support Document and Content management.
- Document and Content Management activities including lifecycle (creation/collection, processing, storage, retrieval, archiving, and destruction).
- Publish, deliver, evaluate and improve.

These activities include the ownership or custody of content such as records, documents, information and other intellectual assets held by the organisation and the governance of information collection, arrangement, storage, maintenance, retrieval, dissemination, archiving and destruction.

The Generally Acceptable Recordkeeping Principles (GARP) can be adopted. These principles are: Accountability, Integrity, Protection, Compliance, Availability, Retention, Disposition, and Transparency.

References:

ARMA International <https://www.arma.org/>

GARP principles: <https://www.arma.org/page/principles>

Data Management Body of Knowledge (DMBOK), Data Management International www.dama.org

Process-level description - GSBPM OP - Knowledge Management

Every cycle of GSBPM generates some content. Complete cycles develop documents on the 8 phases. Operational cycles develop content on the phases 4 to 8. On a full cycle, Mexico's INEGI regulation specifies content as follows:

1. Specify Needs

- Business case, structured need.
- Data availability report.
- Technical and economic viability.
- List of government agencies specifying the need.
- Matrix specified need vs information request.
- List of concepts and domains.
- Data entries.

2. Design

- Conceptual design.
- Design of the production systems and workflows.
- Collection Design.
- Process design including integration, coding, edit, validation, creation of variables, weighing, imputation, estimations, aggregates calculations
- Sample procedure and sample type (if applicable).
- Dissemination plan.

3. Build

- Workflow configuration.
 - Backup of components, systems and software.
4. Collect
- Initial sample load and organisational structure.
 - Training and supervision reports.
 - Non-response report.
 - Relevant logs.
 - Collected Data set.
5. Process
- Collected database image.
 - Changelog.
 - Processed data set.
 - Code used for the weighted process.
 - Vector of weights.
6. Analyze
- Information Product. Including the data set with confidentiality controls, aggregate data, indicators, and metadata.
 - Analysis report.
 - Delivery ready report.
7. Disseminate
- Products and presentations.
 - Structural and Referential Metadata.
 - Log publication for each channel.
 - Marketing strategy.
 - User support log.
8. Evaluate
- Evaluation Report.
 - Action plan.

Relationship

GAMSO defines Document and Content Management that covers the whole organisation. Reference models, policies principles or standards define the actions that need to be followed. GSBPM instances are a subset. The OPs need to highlight the actions that need to be performed at every phase to comply with the specifications established at the GAMSO level. GSBPM operationalises GAMSO Activities at the phase level.

Example

At GAMSO level, an organisation can adopt the GARP principles to define the high-level activities that need to guide the Document and Content Management. At the phase level, GSBPM should make explicit, which documents should be created and stored in the Document and Content Management System.

Case 6. Provider Management

High-level description - GAMSO Activity - Provider Management

GAMSO uses the term Manage Data Suppliers. GSBPM uses the term Provider Management. These activities cover the relationships with data suppliers, which could include public sector and private entities that supply data for statistical activities. These activities include cross-process burden management, as well as topics such as profiling and management of contact information (and thus has particularly close links with statistical business processes that maintain registers. These include:

- Manage data sharing agreements
- Manage data transfer

References:

Paragraph 41 in GAMSO 1.2: Manage Data Suppliers

Process-level description - GSBPM OP - Provider Management

Paragraph 71 in GSBPM 5.1: This sub-process includes the management of the providers involved in the current collection, ensuring that the relationship between the statistical organisation and data providers remains positive, and recording and responding to comments, queries and complaints. Proper communication with reporting units and minimisation of the number of non-respondents contribute significantly to a higher quality of the collected data.”

From Paragraph 14 in GSBPM 5.1: This includes cross-process burden management, as well as topics such as profiling and management of contact information (and thus has particularly close links with statistical business processes that maintain registers).

Relationship

GAMSO defines policies and protocols that govern how Data Suppliers are managed across the organisation. GSBPM OPs operationalise these policies and protocols at the statistical business processing level. There is an explicit dependency between how providers are managed across the organisation and how they are managed at the GSPBM level. For example:

- Where a statistical business process identifies that a data supplier's details or preferences have changed, the enterprise management of this information needs to be updated so the supplier doesn't need to tell the agency multiple times
- A single breach of sound practices by an individual statistical business process risks the reputational standing of the organisation as a whole when it comes to honouring the “contract” with data providers.

Example

GAMSO:

- Provision for management of data supplier in a statistical organisation's strategy for stakeholder management
- Policy to articulate how and when there is a need for a formal agreement with data suppliers
- Policy to minimise the burden of data suppliers when providing data
- Protocol to define the use of data provided by a data supplier
- Provision of digital information and support for data providers.

A concrete example in this space is provided by the UN Handbook of Statistical Organization (2003) which characterises these cross-process activities as Respondent Relations, designed to create a cooperative frame of mind on the part of the intended respondents,

The Handbook emphasises the need to undertake activities in support of the following principles

- A. The purpose of the data collection must be clear and meaningful to the respondent
- B. The statistical agency must be perceived as holding in the strictest confidence all individual records, protecting them from any other party inside or outside Government
- C. The statistical agency must be seen as willing to accommodate respondents, either by providing additional explanations or by accepting legitimate substitutes for a traditional questionnaire
- D. The professionalism and objectivity of the statistical agency, as well as its freedom from political interference, must be established, accepted, and continually advertised
- E. The statistical agency should be perceived as thoughtful and concerned in matters relating to response burden; that is to say, it should be committed to finding means that simplifies the paperwork burden.

These activities are undertaken to ensure (and be seen to ensure) to honour the contract between the respondent and the agency when the respondent provides private data. These activities are central to protecting and promoting the “social license” of the agency to collect, and responsibly process and disseminate data. Further concrete examples include Statistics Canada’s respondent relations management program and the Respondent Advocate approach from Statistics New Zealand GSBPM OP:

Examples of provider management activities include:

- Developing and applying a formal agreement with a provider for the provision of data
- Bringing in test data from providers and assessing it for user needs
- Contacting providers before, during and after a data collection
- Proactively measuring and monitoring provider load and burden, acting on them as needed
- Maintaining a record of all provider contacts to ensure history is available
- Managing provider complaints and feedback for continuous improvement

6. Implementation approaches

The first topic of analysis about the implementation of the models, GAMSO and GSBPM, is to understand what it means to implement the model. In general, we can say that adopting a model does not necessarily mean the incorporation of new activities. It instead means the adoption of the terminology of a model to standardise the identification of concepts and procedures. For instance, in the case of GAMSO, when an NSO decides to incorporate Manage Data Suppliers, it makes explicit the identification of the activities related to the interaction with suppliers. The activities were already there; it is only adopting the identification and formalising the activities. Likewise, adopting GSBPM does not mean the introduction of new activities. It means the adoption of common terminology throughout the production process. Henceforth, for NSOs implementation of HLG-MOS models means the adoption of standard steps and terminology to identify the activities.

One of the common questions of the HLG-MOS model users is: What model should be implemented first, GAMSO or GSBPM? There is no straight answer since either model can be implemented on its own, and both models can be implemented simultaneously. The approach should be based on the needs and resources of each institution.

The approach selected may be based on opportunities. For example, the introduction of either model - or both models - may work best when it is seen to be adding practical value in supporting some other change the organisation is undertaking. If the organisation is seeking to account for time and resources spent on various activities, for example, GAMSO may be a good model to reference, or if it has a general corporate process improvement or quality management initiative. Conversely, if the organisation is

seeking to break down “silos” in statistical production (e.g. across economic and social statistics) then GSBPM may be the model that adds value in the near term.

Henceforth, the adoption approaches can be GAMSO only (high-level), GSBPM only (low-level) and the synergistic adoption of both.

6.1 GAMSO (high-level)

This approach implies the adoption of GAMSO Activities that are usually operationalized in high-level schemas. The implementation relies on the top management of the institution declaring policies, principles or standards that refer to the Corporate Activities defined in the GAMSO Model. The declaration can be one activity at the time or all activities together. A high-level declaration does not guarantee consistent implementation since every process can operationalise the policies, principles, and standards differently.

6.2 GSBPM (low-level)

This approach implies the adoption of the discipline without the existence of high-level policies, principles and standards. In this case, actions, such as metadata management, are adopted at the phase level of GSBPM. Every discipline can be adopted process by process or all processes together.

6.3 Synergistic

Simultaneous adoption of GAMSO and GSBPM is considered a synergistic approach. The identification of GAMSO high-level policies, principles and standards harmonised with the operationalisation of activities at every Phase of the GSBPM can generate positive externalities for the organisation. To adopt both models we can follow the following approaches:

- GAMSO Activities and then the GSBPM OPs.
- GSBPM OPs and then GAMSO Activities.
- Both models together.

The implementation of GSBPM can be performed independently on every statistical program. However, it can also be implemented by defining a standard reference model that applies to the processes of all statistical programs.

7. Conclusion

The development of this note illustrated the convenience to explicitly relate the corporate Activities included in GAMSO with those activities declared in GSBPM. The declaration of high-level GAMSO Activities can provide essential guidance for corporate processes related or unrelated to statistical activities. To achieve a synergistic implementation, the adoption of GSBPM OPs should follow high-level GAMSO guidelines.

To facilitate the adoption of both models, we identify the following opportunities:

- Adopt consistent terminology in GAMSO and GSBPM by referring to a standard glossary.
- Define all statistically related GAMSO Activities to a correspondent GSBPM OP.
- Explore if other knowledge areas such as Geographic Information Systems are relevant for the production of statistical information.
- Include, at the phase level, OPs descriptions on the GSBPM paper.
- Develop an explicit relationship of the GAMSO Activities and GSBPM OPs with the Data Management knowledge areas defined in the Data Management Body of Knowledge (DMBOK).

Annex A. Historical perspective of the relationship between GSBPM and GAMSO

By Steven Vale, UNECE, May 2019

From the mid-1980s until 2013 the UNECE ran a Steering Group on Statistical Metadata, known as “METIS”. In 2004, the METIS group started to create a Common Metadata Framework (CMF), including the development of “Common/generic models for statistical metadata”.

A Workshop in Vienna in 2007 looked at existing national models and agreed to develop a “Statistical Survey Life Cycle Model” based on the approach used in New Zealand. This was soon re-named the Generic Statistical Business Process Model (GSBPM). Versions 1-3 were discussed within the METIS group and were further refined during 2008. A Workshop in Lisbon on March 2009 agreed that version 4.0 was ready for use. It was released on the public “METIS Wiki” the following month.

As the GSBPM was developed in the context of statistical metadata, it only needed to cover statistical production. The METIS group recognised that other activities existed in statistical organisations. As most of these activities were not important for METIS purposes, they were just listed as “over-arching processes”. However, two “over-arching processes” (metadata management and quality management) were included in the GSBPM picture, and a few paragraphs were added about them at the end of the GSBPM document. The text on metadata management was mostly copied from Part A of the CMF “Statistical Metadata in a Corporate Context”.

The intention was to come back to the issue of “over-arching processes”, and model them in more detail in future versions of the GSBPM.

In 2014-15, the Statistical Network was developing a “Business Activity Model” (BAM). This model included a “production” part, which was similar to the GSBPM. The other parts of the BAM seemed to cover the GSBPM over-arching processes. The UNECE High-Level Group for the Modernisation of Official Statistics (HLG-MOS) had, by that time, taken responsibility for the GSBPM, and saw an opportunity for a “quick-win” by using the BAM as a basis to extend the GSBPM to better cover the “over-arching processes”.

An HLG-MOS task team was set up in mid-2014 to use the BAM as a basis for creating a generic extension to the GSBPM. Various names were considered, including “Generic Business Activity Model”, “Generic Statistical Activity Model” and “GSBPM+”, before the name “Generic Activity Model for Statistical Organisations” (GAMSO) was chosen as a compromise solution. Version 0.4 of the new model (written before the name was finalized), described the purpose of the model as follows:

“The [name] describes and defines the activities that take place within a “typical” statistical organisation. It extends and complements the Generic Statistical Business Process Model (GSBPM) by adding the activities needed to support statistical production. When the GSBPM was developed, such activities were referred to as “over-arching processes”, and were listed, but not elaborated in any real detail. Over the years, there have been several calls to expand the GSBPM to better cover these activities. The [name] was therefore developed to meet these needs.”

“Like the GSBPM, the [name] aims to provide a common vocabulary and framework to support international collaboration activities, particularly in the field of modernisation. However, as the GSBPM is becoming increasingly used for cost and resource management within organisations, the [name] is also expected to fill such roles when there is a need to consider a wider context than just the activities directly related to the production of statistics.”

Thus, the intention was that the GAMSO would replace the “over-arching processes” of the GSBPM, and the GSBPM would be the statistical production part of the GAMSO. The intended relationship between GAMSO v1.0 and GSBPM v5.0 is shown in Figure 6 below.

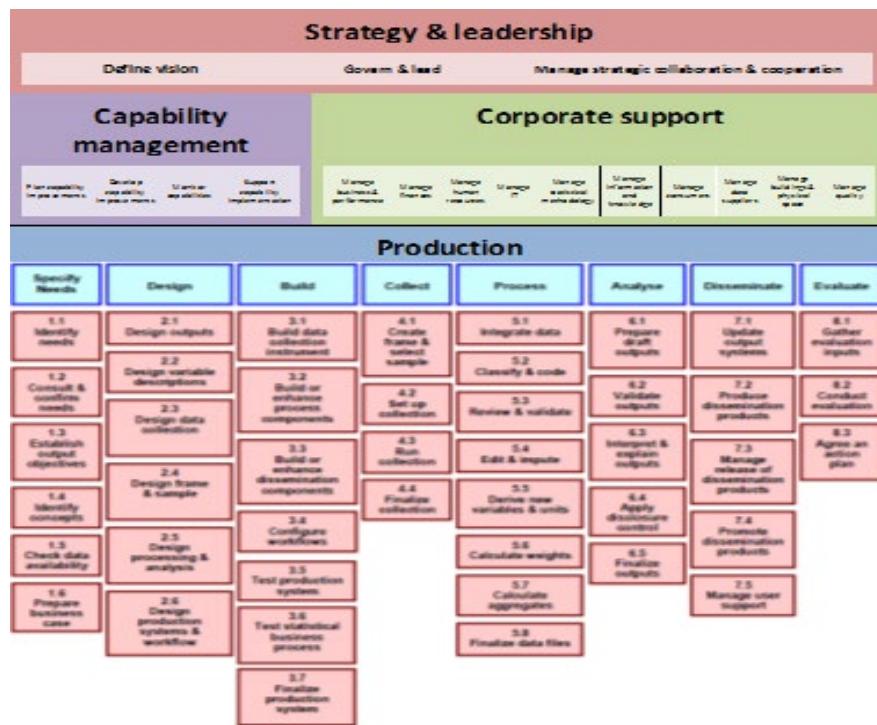


Figure 6: GAMSO v1.0 and GSBPM v5.0 – Combined view

GAMSO v1.0 was presented to and approved by the HLG-MOS on this basis, and was released on March 2015. It was decided not to revise the GSBPM documentation to remove the over-arching processes at this point because version 5.0 had been released at the end of 2013 and the HLG-MOS had agreed on a 5-year revision cycle. However, training materials and presentations were adjusted to show how the GAMSO and the GSBPM fitted together. The plan was that the models would be formally reviewed and revised together in 2018, at which point they would be made entirely consistent.

Annex B. List of abbreviations

- DAMA - [Data Management Association International](#)
DDI - [Data Documentation Initiative](#)
DMBOK - [Data Management Body of Knowledge](#)
GAMSO - [Generic Activity Model for Statistical Organizations](#)
GSBPM - [Generic Statistical Business Process Model](#)
GSIM - [Generic Statistical Information Model](#)
HLG-MOS - [High-level Group for the Modernisation of Official Statistics](#)
OP - Overarching Process.
SDMX - [Statistical Data and Metadata eXchange](#)
UNECE - [United Nations Economic Commission for Europe](#)