

# Navigating the application of Modernisation Frameworks when using Commercial off the Shelf products.

## Introduction

The ABS has been a key contributor to the development and maturation of the current Modernisation Frameworks for Statistical Organisations.

Specifically, the ABS was a key member of the Stats Network responsible for the development of the Generic Statistical Information Model (GSIM), as well as maturing the Generic Statistical Business Process Model (GSBPM) to version 5.0.

The ABS began a major transformation program in 2013, designed to standardise and modernise our core processes and systems. The architecture developed to guide this transformation was fundamentally based on the current Modernisation Frameworks.

This paper will discuss how the ABS used these Modernisation Frameworks to guide the implementation of automated, metadata driven business processes, and the challenges faced when utilising a Commercial of the Shelf product supported by open standards.

## Frameworks to drive transformation

The ABS, like many other National Statistical Organisations (NSO's), found itself with poorly documented, inconsistent, aged and siloed processes and systems.

A large part of the problem stemmed from having different terminology used from one Statistical Program to another, and having bespoke system and processes built based on that siloed terminology. This was of course exacerbated with terminology differences between different NSO's.

The [Statistical Network](#) comprised seven NSO's (ABS, ONS, Stats Norway, Stats Sweden, INSEE, Stats Canada and Stats NZ). Work began to develop a framework which would describe in a consistent way, the information objects used in the production of official Statistics, and the relationships inherent between these objects.

The [Generic Statistical Information Model \(GSIM\)](#) was developed as a reference framework of information objects, which enables generic descriptions of the definition, management and use of data and metadata throughout the statistical production process. This framework is now owned and governed by the UNECE Statistics.

This framework formed the core artefact of the ABS Information Architecture layer within our Enterprise Architecture. The ABS developed the ABS Information Model (AIM), whilst based on GSIM, goes beyond the scope of GSIM as a reference model. AIM specifies the logical and physical modelling required supporting our Metadata Registry & Repository (MRR) and Enterprise Data Management Environment (EDME).

The ABS has also "de-scoped" some aspects of GSIM (e.g. Conceptual Domains, Category as a class separate to Concept) because:

1. business feedback indicated that populating these separately on a consistent basis would add considerable complexity and effort for ABS staff without commensurate return, and
2. ABS could “reconstitute” these classes based on other classes and relationships in AIM if we ever needed to map metadata to GSIM.

The ABS also utilised the common statistical framework for describing and categorising activities undertaken in the production of official statistics. The [Generic Statistical Business Process Model \(GSBPM\)](#) describes and defines the set of business processes needed to produce official statistics.

The GSBPM includes activities and processes which the ABS is seeking to make redundant through our transformation program. The GSBPM however needs to remain relevant to all NSO’s, including those who are still developing their statistical capabilities.

The ABS has developed its own variant on the GSBPM to include more advanced activities, and remove those which will no longer be required in the post transformed ABS. The Statistical Production Activity Model (SPAM) is the key artefact within the ABS Business Architecture layer, and is used as the key framework for defining the ABS’s reusable Business Processes. At a high level, SPAM helps to articulate the “to be” target state for ABS business process design, making it more specific and “normative” than GSBPM, rather than simply “descriptive”.

It must be noted that the ABS has detailed mappings back to the GSBPM to aid with international collaboration efforts.

## Commercial of the Shelf Business Process Automation Platforms

Gartner defines Business Processes as “The coordination of the behaviour of people, systems and things<sup>1</sup> to produce specific business outcomes.”

Business Process automation platforms minimally include:

- a graphical business process and/or rule modelling capability,
- a process registry/repository to handle the modelling metadata,
- a process execution engine and a state management engine or rule engine (or both).

These platforms are supported by open standards such as:

- [BPMN 2.0](#)

The Business Process Modelling Notation (BPMN) is a graphical notation that depicts the steps and the end to end flow of a business process. The notation has been specifically designed to coordinate the sequence of processes and the messages that flow between different process participants in a related set of activities.

It should be noted that the [BPMN standard influenced the original business group modelling in GSIM](#). The standard was simplified and included as part of GSIM to support NSO’s which did not have, nor required a full COTS BPMN product.

---

<sup>1</sup> "Things" in this context refers to devices that are part of the Internet of Things (IoT).

Where an NSO was implementing a BPM platform, the open standards of BPMN were to be utilised over GSIM.

- [BPEL](#)

BPEL is an XML-based language for describing a business process in which most of the tasks represent interactions between the process and external Web services. The BPEL process itself is represented as a Web service, and is realised by a BPEL engine which executes the process description.

- [WS-HumanTask](#)

WS-HumanTask enables the integration of tasks performed by human beings in service-oriented applications. It provides a notation, state diagram and API for human tasks, as well as a coordination protocol that allows interaction with human tasks in a more service-oriented fashion and at the same time controls tasks autonomy.

The ABS acquired a Commercial off-the-shelf (COTS) Business Process automation tool to document, manage and execute its automated business processes.

[ActiveVOS](#) (Active Visualisation Orchestration System) is a BPM product, acquired by Informatica in 2013 after five years of independent development, which:

- is a service orientated business process automation platform.
- executes BPMN models directly on a high-performance BPEL engine that runs on any standards-based Java Enterprise Edition server.
- supports open standards such as BPMN 2.0, BPEL and WS-Human-Task.

With ActiveVOS supporting these open standards, it became unnecessary to implement the full model for Business Processes as described within GSIM.

## **The ABS Process Architecture and the Business Process Information Model**

The ABS Process Architecture supports the standardisation and re-use of Business Processes.

This was never going to be achieved however by attempting to standardise processes from the top down (i.e. standardising the entire end to end processes). Whilst the lower level activities and supporting processes may be the same, different Statistical Programs perform these activities in different orders and/or may not even undertake all the activities as other Statistical Programs.

This is often for valid statistical design (e.g. methodological) reasons, whilst other times it is more “we’ve always done it that way” and/or “it offers marginal benefits locally but not sufficient to warrant the cost and complexity imposed by supporting a standard process corporately”.

Ideally a combination of top down and bottom up helps separate “necessary variants” from “unnecessary variants”; however this is more a corporate governance concern.

The Process Architecture was developed to support both approaches equally, providing the organisation with flexibility by standardising processes “building blocks” for activities described by the Statistical Production Activity Model (SPAM) at level three (or below). These “building blocks” can then be utilised within Workflows for a Statistical Program in any order which meets their



The other major change from GSIM was the replacement of the:

- Process Input Specification and Process Output Specification objects with the Parameter Set Definition.
- Process Input and Process Output objects with the Parameter Set.

This allows Parameter Set Definitions to be defined in their own right where Input Specifications and Output Specifications at a conceptual level did not have an explicit grouping class. So they could be seen as a formalisation of a class that was only (potentially) implicit in GSIM. Formalisation supports things like hierarchical design and registration as a form of “standard contract” for working with a standard business process etc.

It also more formally “types” actual inputs and outputs back to the specification. So this was a “necessary elaboration” required compared with GSIM. In terms of actual inputs and outputs, the elaboration more neatly separates which are local defaults (design time) and which are local run time choices.

The Parameter Sets contain two distinct types of information:

- Entity Content – this is reference to other metadata registered in the MRR which will be used within the Business Process
- Name Value Pair Type – this is to capture all other information required by a Business Process, and includes Boolean, integer, strings, time values etc.

This enables the capture and storage of all required configuration information for Statistical Programs Workflows, and the re-use of standardised Business Processes.

The objects within the Business Process Information Model are defined as:

- Business Process  
The Business Process is the registration of a standard, re-usable and configurable process supporting a defined business activity. This business activity is described the Statistical Production Activity Model (SPAM) typically at level three or below.
- Parameter Set Definition  
The Parameter Set Definition specifies the types of Process Inputs and Process Outputs that the Business Process will either consume (the configuration) or produce (the record of output).
- Workflow  
A Workflow is a process for a Statistical Program which undertakes a higher level business function. Typically, Workflows will orchestrate together many reusable Business Processes.
- Process Step  
The Process Step represents the re-use of a Business Process within a Statistical Program’s Workflow. The Process Step binds the Parameter Set (configuration) for the re-use of the Business Process.
- Parameter Set  
The Parameter Set is used to capture and store the design and runtime time Process Inputs for a Statistical Program’s intended re-use of the Business Process. The Parameter Set is described by

the Parameter Set Definition which is related to the Business Process that the Process Step is calling. The Parameter Set is also used to store a record of the Process Outputs once executed.

- Workflow Instance  
The Workflow Instance is the registration in the MRR of the execution of a Workflow for a Statistical Program.
- Process Step Instance  
The Process Step Instance is the registration in the MRR of the execution of a Process Step within a Workflow Instance for a Statistical Program. The Process Step Instance binds the Parameter Set for the actual re-use of the related Business Process.

The remaining objects retain their GSIM definitions within the ABS Business Process Information Model.

- [Statistical Program](#)  
A set of activities, which may be repeated, that describes the purpose and context of a set of Business Process within the context of the relevant Statistical Program Cycles.
- [Statistical Program Cycle](#)  
A set of activities to investigate characteristics of a given Population for a particular reference period.
- [Statistical Program Design](#)  
The specification of the resources required, processes used and description of relevant methodological information about the set of activities undertaken to investigate characteristics of a given Population.
- [Statistical Support Program](#)  
A program which is not related to the post-design cyclic production of statistical products, but is necessary to support cyclical production.

The use of these objects within the ABS Business Process Information Model is described in the below diagram (which is also stepped through in the accompanying PowerPoint presentation):

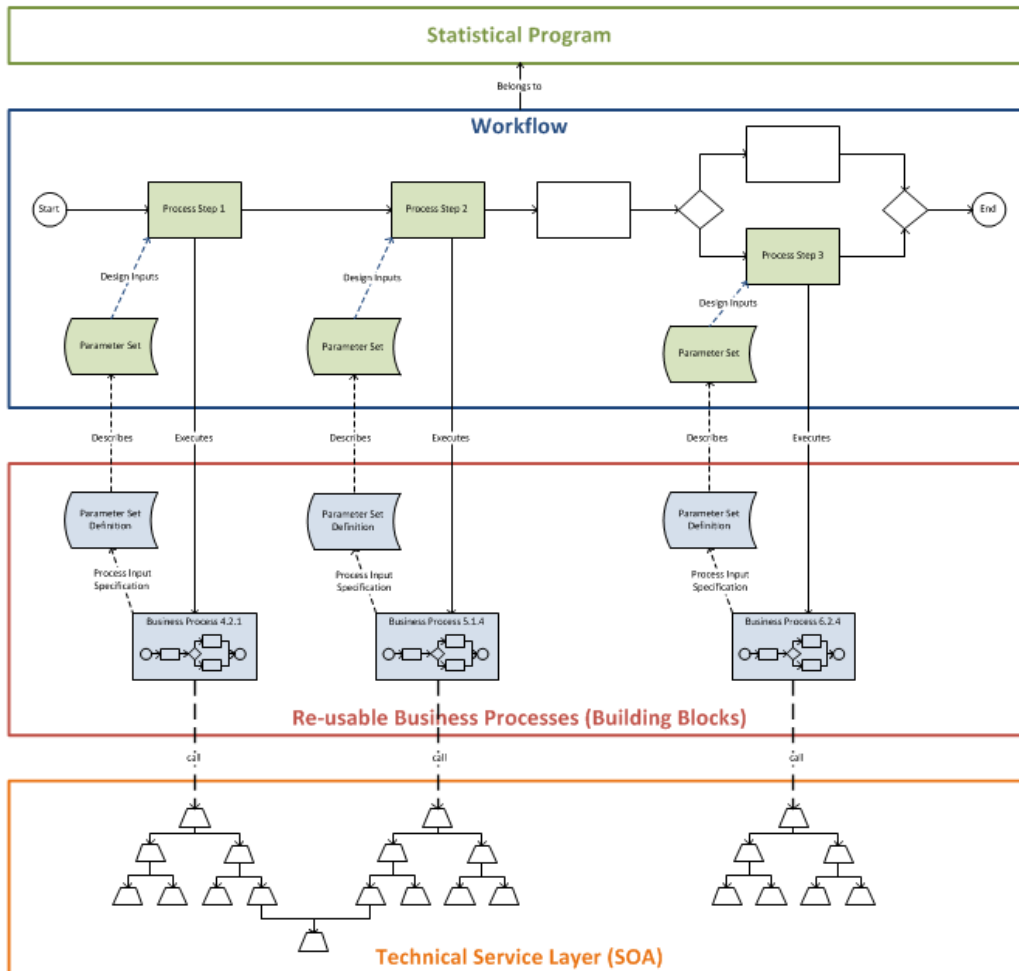


Figure 2 - Workflow composition

## Conclusion

The ABS in acquiring a COTS Business Process automation product faced the challenge of utilising the Modernisation Frameworks whilst not duplicating the open standards natively supported by that product.

This meant examining the core requirements for registering Business Processes within our Metadata Repository, and then modifying the GSIM Business Group to meet these requirements.

The ABS has successfully implemented a minimal Business Process Information Model supporting the registration of Business Processes and Workflows within our Metadata Registry and Repository (MRR), as well as capturing the process configurations with the use of parameter sets.

This has ensured full discoverability and support for standardised, re-usable Business Processes by utilising our core foundation infrastructure (namely the MRR, whilst taking advantage of all the benefits that come with a COTS BPM product.

Agencies that apply GSIM but do not have a BPM product based on open standards may find that the full “richness” of the original modelling in GSIM continues to add value for them. Agencies that do use such a BPM product may find, similarly to the ABS, that GSIM can be applied in a way that complements, rather than duplicates, BPM capabilities.