

**UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN UNION (EUROSTAT)**

**ORGANISATION FOR ECONOMIC COOPERATION
AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

Work Session on Statistical Metadata

(Geneva, Switzerland 6-8 May 2013)

Topic (iii): Metadata in the statistical business process

METADATA FLOWS IN THE GSBPM

Working Paper

Prepared by the Informal Task Force on Metadata Flows

I. Introduction

1. The Workshop on Statistical Metadata 2011 “Implementing the GSBPM and Combining Metadata Standards” included a session on “(v) Metadata flows within the GSBPM”. This session considered what kind of and how statistical metadata are created, managed and used within and between the different phases of the Generic Statistical Business Process Model (GSBPM); A group discussion, where the participants worked on a proposal prepared for the theme, was held.
2. Participants considered this work very important and one conclusion of the workshop was that “a new informal group should be set up with a focus on data and metadata flows within the GSBPM. This group would work on creating a systematic description of processes, metadata objects and flows, and would expand the documentation produced by Statistics Portugal for topic (v).”
3. Meanwhile, work on a Generic Statistical Information Model (GSIM) was following its path with increasing importance and the risk of duplication of effort with the work of such an informal group was high. It was decided by the METIS Steering Group to wait for GSIM 0.8, and then decide how to take forward the “informal group” on metadata flows within GSBPM.
4. In October 2012 it was decided by the METIS Steering Group to create an informal task force (TF) on metadata flows; the work of this task force would take place 'virtually' via teleconferencing, Webex and/or wiki communication, with secretariat and technical support to be provided by UNECE.
5. In the scope of the METIS Steering Group it was proposed as a basis for the work of the TF to determine which information objects form the input and output for each GSBPM sub-process. While some refinements to the GSBPM might need to be considered in the light of the development of GSIM, the proposal would be helpful for achieving this goal.
6. Meanwhile, GSIM 1.0 was made available on the 21st of December 2012, and the TF work was done taking into account this last version of GSIM.

II. Composition of the Task Force

Alistair Hamilton (ABS)
Anne Gro Hustoft (SSB)
David Brackfield (OECD)
Ebbo Petrikovits (CZSO)
Eric Rodriguez (Mexico)
Hakan Linden (Eurostat)
Isabel Morgado – TF coordinator (Statistics Portugal)
Jenny Linnerud (SSB)
John Csaki (ABS)
Justin Lynch (ABS)
Rute Costa (Eurostat)
Thomas Nyberg (SCB)
Tim Dunstan (SC)

With the support of Fiona Willis-Núñez (UNECE).

III. Objectives of the Work

7. Keeping in mind that GSIM intends to be a conceptual reference model for statistical information, while GSBPM intends to be a reference framework for statistical organizations to agree on standard terminology to describe and define the business processes needed to produce official statistics and also to enable harmonizing statistical computing infrastructures and facilitating the sharing of software components, the main goals of this work are:
 - Explore the relationship between GSIM and GSBPM and establish the flow of the information objects across the sub-process and the role(s) of each information object in each sub-process;
 - Make this relationship understandable to the different populations that deal with these models (managers, specialists), mainly for managers in this phase;
 - Test if the work done on the GSIM model until now fits the purpose for which it was conceived;
 - Evaluate if the GSBPM model needs some refinements;
 - Contribute to transform these two models into true standards;
 - Contribute to harmonization of terminology across statistical organizations;
8. To achieve these goals, some tasks need to be carried out to add value to the previous work done on GSIM and GSBPM, and determine the concrete objectives of this task force, that are based on GSBPM (which establishes standard sub-processes of the statistical business process), and on the GSIM 1.0 (which lists and defines information objects supporting statistical business process, most of them metadata objects):
 - Establish, for each sub-process of GSBPM, which GSIM 1.0 information objects (IOs) are input and which are output;
 - Specify process steps for processing input IOs and production of output IOs;
 - Identify information objects, which are neither input nor output, but support sub-processes;
 - To present the work done at the 2013 METIS Work Session.

IV. Methodology of the Work

9. Two templates were used for the work:
 - (a) Detail of each sub-process:

10. List the information objects (IO) that are input and output to the sub-process and detail the operations over each IO during the process step. Assuming that (CRUD – Create, Read Update, Delete) is more IT terminology, the team decided use terms in more natural language because these are more comprehensible and flexible in their utilization; so the terms used were:

Create – When an IO is brought into existence;

Use – Reuse an already existing IO;

Consult – To seek an answer from someone or from reference material;

Analyse – Read and make a judgment on the IO read;

Update – Make some change in an already existing IO;

Delete – Remove the IO;

Sub-process X.Y - Title of sub-process X.Y

<i>Process Input</i>	<i>Process Step</i>	<i>Process Output</i>
<p>Transformable input:</p> <p>Parameter input:</p> <p>Process support:</p> <p>and specify the sub-process A.B that they come from</p>	<p>Indicate what happens to each of the Process Inputs and the Process Outputs in this phase: (Create/Use/Consult/Analyse/Update/Delete)</p> <p>Processed input:</p> <p>Production of outputs:</p>	<p>Transformed output:</p> <p>Process metric:</p>
	<p>Supporting Guidelines/Documentation</p> <p>► Guideline 1</p> <p>► Guideline 2</p>	<p>Feedback</p> <p>► Request update of Input x from sub-process A.B</p>

Table 1 – Sub-process table

11. The feedback section in the table is used to signify the consequence of this sub-process in other sub-processes and if some action must be taken, even in previous sub-processes, establishing in this way an iteration in the sub-processes of the phase we are in or in other phases.
12. The GSIM 1.0 Specification document contains about 150 information objects, at very different levels of detail.
13. The work began with all the GSIM 1.0 information objects but, doing this work with so many information objects was not easy, so the team decided to restrict the number of information objects to those of a higher level, working with the 21 information objects described in Figure 3 of the Communication document, with minor changes (see Table 2 below).
14. By including only 21 information objects, the result should be more comprehensible by managers and statisticians in general.
- (b) Overall view of the IOs used in the sub-processes:
15. The colours used to stress the group of several IOs are those used in the GSIM 1.0 specification document.

Information object	Phase		
	Sub-Process	Sub-Process	Sub-Process
Acquisition Activity			
Assessment			
Business Case			
Dissemination Activity			
Production Activity			
Statistical Activity			
Statistical Need			
Statistical Program			
Statistical Program Design			
Process Input			
Process Output			
Process Step			
Classification			
Concept			
Population			
Unit			
Variable			
Information Resource			
Data Set			
Data Structure			
Product			

Table 2 – Restricted set of information objects proposed (21)

16. In order to harmonize the work done and to facilitate this job for each subgroup, a top-down approach to the 9 phases was used for these few information objects; This high level work forced some additions to the list of IOs to use.
17. All the sub-processes were done/redone with these few information objects, and therefore harmonized.
18. Difficulties didn't end with this approach, because different persons have different understandings about what information objects are, although their definitions are provided in the GSIM Specification - Annex C. Furthermore, the terminology used by people in different statistical institutions is different.
19. A third approach was tried: a set of assumptions at the phase level were proposed, in order to help everyone to think in the same way about the same information objects.
20. Accepting these assumptions at the phase level, the work was adjusted at a sub-process level, following the sequence of phases in GSBPM.
21. With the IOs identified for all the sub-processes, it was found that some IOs, which weren't considered initially, were necessary and some of the IOs considered weren't used; table 3 (below) was born with the IOs actually used, for the 9 phases.
22. A table was prepared with all the sub-processes and the IOs used, with inclusion of the operations performed on each IO in each sub-process. This table gave the team an overall view of the statistical

business process and the role of each IO in each sub-process. This table was very useful and brought several benefits:

- Harmonization of the several sub-process tables (table 1) done by each subgroup;
- Identification of the IOs actually used and needed for the work;

23. The IOs used are listed in the next table (table 3); actually the team worked with 25 high level information objects.

Information object	Phase		
	Sub-Process	Sub-Process	Sub-Process
Assessment			
Statistical Need			
Change Definition			
Business Case			
Statistical Program Design			
Process Step Design			
Process Method			
Rule			
Instrument			
Process Input			
Process Output			
Process Metric			
Process Step			
Process control			
Process Step Execution Record			
Classification			
Concept			
Population			
Unit			
Variable			
Information Resource			
Data Set			
Data Structure			
Product			
Provision Agreement			

Table 3- Restricted set of information objects actually used

The complete table can be consulted in Annex I to this text.

V. Examples of the Work Done

24. In this topic, we give some examples of sub-processes worked on:

Phase 2 – Design

2.1 – Design Outputs

“This sub-process contains the detailed design of the statistical outputs to be produced, including the related development work and preparation of the systems and tools used in phase 7 (Disseminate). Outputs should be designed, wherever possible, to follow existing standards, so inputs to this process may include metadata from similar or previous collections, international standards, and information about practices in other statistical organizations from sub-process 1.1 (Determine need for information).”

Process Input	Process Step (Create/Analyse/Consult/Use/Update/Delete)	Process Output
<p>Transformable input:</p> <ul style="list-style-type: none"> Partial, high level <i>Statistical Program Design</i> setting out agreed features based on <i>Business Case</i> (from 1.6) This forms part of the partial high level <i>Statistical Program Design</i>. <p>Parameter input:</p> <ul style="list-style-type: none"> n/a <p>Process support input:</p> <ul style="list-style-type: none"> Partial, high level <i>Statistical Program Design</i> setting out agreed features based on <i>Business Case</i> (from 1.6) <i>Statistical Need</i> (1.1) <i>Change Definition</i> (1.2) <i>Assessment</i> (1.3) <i>Assessment</i> (Gap Analysis) (1.5) <i>Units</i> (1.3) <i>Populations</i> (1.3) <i>Concepts</i> (1.4) <i>Variables</i> (1.4) 	<p>Input:</p> <p>Consult the partial <i>Statistical Program Design</i> and the <i>Business Case</i> to understand the agreed output objectives and the design features and constraints which will apply from a business perspective when addressing these objectives.</p> <p>Consult dissemination policies, guidelines and toolsets to understand available options for dissemination formats and channels.</p> <p>Consult <i>Statistical Need</i> (1.1), <i>Change Definition</i> (1.2), <i>Assessment</i> (1.3) and <i>Gap Analysis</i> (1.5) for additional detail.</p> <ul style="list-style-type: none"> Consult <i>Units</i> (1.3), <i>Populations</i> (1.3), <i>Concepts</i> (1.4) and <i>Variables</i> (1.4) to better understand user objectives related to the output. Consult agency standards in regard to defining output <i>Variables</i> and <i>Classifications</i>. <p>Output:</p> <p>Update design specification of the <i>Information Resource</i> to be produced via the Statistical Business Process. The resultant specification must be to a sufficient level of detail to identify what content will be made available through <i>Products</i> and <i>Dissemination Services</i>.</p> <p>Updating the specification of the <i>Information Resource</i> may lead to updating the specification of <i>Units</i>, <i>Populations</i>, <i>Concepts</i>, <i>Variables</i> and <i>Classifications</i> which will actually be represented in the output.</p> <p>Design the <i>Products</i> and <i>Dissemination Services</i> which will be made available to disseminate outputs from the Statistical Business Process. This includes a broad description of the content which will be available via <i>Products</i> and <i>Dissemination Services</i> (eg what cross tabulations will be available for key output <i>Variables</i> to what level of detail to what level of statistical quality for what (sub) <i>Populations</i>).</p>	<p>Transformed output:</p> <ul style="list-style-type: none"> Updated <i>Information Resource</i> (design only) <ul style="list-style-type: none"> The design of the <i>Information Resource</i> to be produced via the Statistical Business Process. <i>Products</i> (design only) <i>Dissemination Services</i> (design only) <ul style="list-style-type: none"> Includes specification of which <i>Dissemination Services</i> will be used and (broadly) the content to be made available through each. Updated <i>Statistical Program Design</i> <ul style="list-style-type: none"> Particular focus on <i>Dissemination Design</i>. Includes summary of <i>Process Steps</i> to be undertaken in Phase 7. <p>Secondary</p> <p>The process of designing outputs may lead to updates to specifications of <i>Units</i>, <i>Populations</i>, <i>Concepts</i>, <i>Variables</i> and <i>Classifications</i> which will actually be represented in the output.</p> <p>This is not a feedback loop to 1.3 and 1.4 - the user specifications in regard to <i>Units</i>, <i>Populations</i>, <i>Concepts</i> and <i>Variables</i> should not be updated. The specification of what will actually be produced as outputs, however, may need to be updated. These updates will then feed into later sub-processes in the Design phase.</p> <p>Process Metric:</p> <ul style="list-style-type: none"> n/a

Notes on sub-process 2.1 – Design outputs

Assumed progressive definition of the *Statistical Program Design* during GSBPM Phase 2

Initiation of GSBPM Phase 2 (new design work) is assumed to occur when an agency decides to implement a proposition set out in a (original or revised) *Business Case* (1.6).

An agency deciding to implement a *Business Case* would imply that

- one or more existing *Statistical Programs* will have their *Statistical Program Design* updated and/or
- one or more new *Statistical Programs*, with new *Statistical Program Designs*, will be initiated

GSBPM's point of reference is a Statistical Business Process. From a GSIM perspective, each Statistical Business Process will be a *Statistical Program* with an associated (potentially versioned) *Statistical Program Design*.

While a single *Business Case* may lead to several different Statistical Business Processes being initiated or revised, when we describe Phase 2 we are talking about it in the context of a particular Statistical Business Process and therefore in the context of a particular *Statistical Program*.

It is assumed that partial high level outlines of new *Statistical Program Designs*, and/or of new versions of existing *Statistical Program Designs*, can be considered as initiated at the time the *Business Case* is agreed. For conciseness, each case of a completely new *Statistical Program Design* and each case of a new version of an existing *Statistical Program Design* is simply referred to as a "new" *Statistical Program Design* from here on.

Prior to the Design phase commencing in earnest, the new *Statistical Program Design* can be considered as being "prepopulated" with high level design features that were specified and agreed as part of the *Business Case*. The Design phase leads to the *Statistical Program Design* becoming fully developed.

It is further assumed that the *Business Case*, and therefore the initial outline of the *Statistical Program Design*, includes a high level description of the new *Information Resource* which will be produced as a result of the Statistical Business Process.

Assumed "rolling-up" of outputs from earlier sub-processes in Phase 1 by the *Business Case*.

It is assumed that *Business Case*, the primary output from 1.6, recognises and "rolls up" the outputs from earlier sub-processes in Phase 1. The *Business Case* encapsulates how the Design phase should proceed based on those earlier outputs.

One set of examples are *Statistical Need* (1.1), *Change Definition* (1.2), *Assessment* (1.3) and *Gap Analysis* (1.5). While undertaking sub-processes in the Design phase there may be direct reference to those earlier outputs - seeking additional information or clarification. Primary reference, however, will be to the *Business Case* which set out the broad way forward based on those earlier inputs.

Similarly 1.3 produced outputs related to *Units* and *Populations* from a user needs perspective and 1.4 produced outputs related to *Concepts* and *Variables* from a user perspective "which may not align with existing standards". It is currently assumed the *Business Case* will provide high level direction on the extent to which the Design is expected to maintain alignment with existing standards, and support comparability with existing sources, versus the extent to which it should support "new and different" output definitions that more closely correspond with the needs as set out by the users.

Assumed scope of sub-process 2.1

The description of the sub-process in GSBPM V4.0 makes it clear that 2.1 includes design of output products (eg publications, spreadsheets, on line databases). This could be considered design of *Products* and *Dissemination Services* in terms of GSIM V1.0.

What is not stated so explicitly is the extent to which this sub-process includes adding more detail in terms of the design of the *Information Resource* the Statistical Business Process is going to produce. As GSIM V1.0 states, a *Dissemination Service* exposes the *Information Resource* – so to design the *Dissemination Service* in this sub-process we need to have an idea of the *Information Resource* which will be produced for output purposes. For example, what essential cross tabulations, to what level of detail and statistical quality, are expected to be possible in terms of the "output" *Variables* to be produced? (While user needs/objectives in this regard were established in 1.3, 2.1 addresses the extent to which the outputs from the Statistical Business Process are expected to meet these needs).

User objectives in terms of *Units* and *Populations* were established in 1.3. User objectives in terms of *Concepts* and *Variables* were established in 1.4. Factors such as viability of collecting input concepts, cost, alignment with standards and harmonisation with existing sources may mean, however, that the design of the *Information Resource* to be produced may not correspond directly to the outputs from 1.3

and 1.4. The *Business Case* is assumed to contain some degree of strategic decision making and direction in this regard. 2.1 is then assumed to further crystallise the design of the *Information Resource* to be produced and therefore (broadly) the *Units*, *Populations*, *Concepts*, and *Variables* which will actually be represented in the output.

Phase 3 – Build

3.1 Build data collection instrument

“This sub-process describes the activities to build the collection instruments to be used during the phase 4 (Collect). The collection instrument is generated or built based on the design specifications created during phase 2 (Design). A collection may use one or more modes to receive the data, e.g. personal or telephone interviews; paper, electronic or web questionnaires; SDMX hubs. Collection instruments may also be data extraction routines used to gather data from existing statistical or administrative data sets. This sub-process also includes preparing and testing the contents and functioning of that instrument (e.g. testing the questions in a questionnaire). It is recommended to consider the direct connection of collection instruments to the statistical metadata system, so that metadata can be more easily captured in the collection phase. Connection of metadata and data at the point of capture can save work in later phases. Capturing the metrics of data collection (paradata) is also an important consideration in this sub-process.”

Process Input	Process Step (Create/Analyse/Consult/Use/Update/Delete)	Process Output
Transformable input: <ul style="list-style-type: none"> ▶ n/a Parameter Input: <ul style="list-style-type: none"> ▶ n/a Process Support Input: <ul style="list-style-type: none"> ▶ Completed <i>Statistical Program Design</i> from 2.6 ▶ Documented set of <i>Process Step</i> designs for Phase 4 to Phase 8 from 2.6 ▶ List of new system solutions (if any) that need to be built before an <i>Acquisition Activity</i> can be conducted in accordance with the <i>Statistical Program Design</i> from 2.6 ▶ <i>Information Resource</i> with <i>Acquisition Design</i> and <i>Instrument implementation</i> from previous cycles of existing statistics 	Input <ul style="list-style-type: none"> Consult and analyse <ul style="list-style-type: none"> ✓ <i>Information Resource</i> ✓ <i>Statistical Program Design</i> ✓ <i>Process Step Design for phase 4</i> Output <ul style="list-style-type: none"> Create/update <i>Instrument Implementation</i> Update <i>Information Resource</i> Create <i>Process Metric</i> Supporting Guidelines/Documentation <ul style="list-style-type: none"> ▶ National guidelines for questionnaire design ▶ International guidelines for questionnaire design 	Transformed Output: <ul style="list-style-type: none"> ✓ <i>Information Resource</i> with new <i>Instrument Implementation</i> Process Metric: <ul style="list-style-type: none"> ▶ n/a
		Feedback <ul style="list-style-type: none"> ▶ Request update in 2.6 as appropriate

Phase 4 – Collect

4.1 - Select sample

“This sub-process establishes the frame and selects the sample for this iteration of the collection process, as specified in sub-process 2.4 (Design frame and sample methodology). It also includes the coordination of samples between instances of the same statistical business process (for example to manage overlap or rotation), and between different processes using a common frame or register (for example to manage overlap or to spread response burden). Quality assurance, approval and maintenance of the frame and the selected sample are also undertaken in this sub-process, though maintenance of underlying registers, from which frames for several statistical business processes are drawn, is treated as a separate business process. The sampling aspect of this sub-process is not usually relevant for processes based entirely on the use of pre-existing data sources (e.g. administrative data) as such processes generally create frames from the available data and then follow a census approach.”

Process input	Process step	Process output
<p>Transformable input:</p> <ul style="list-style-type: none"> ▶ n/a <p>Parameter Input:</p> <ul style="list-style-type: none"> ▶ Sampling criteria, (2.4); ▶ Sampling methodology, (2.4); <p>Process support:</p> <ul style="list-style-type: none"> ▶ <i>Statistical Program Design</i> <ul style="list-style-type: none"> ✓ Sample Design ✓ Process steps ✓ Business Services (SW Applications) <ul style="list-style-type: none"> • Application to select the frame • Application to select the sample ▶ <i>Information resource</i> <ul style="list-style-type: none"> ✓ Metadata database ✓ Units register ✓ Sampling Frame ▶ <i>Population</i> 	<p>(Create, Use, Consult, Analyze, Update, Delete)</p> <p>▶ Input:</p> <ul style="list-style-type: none"> Consult <i>Statistical Program Design</i> <ul style="list-style-type: none"> ✓ Sample Design ✓ Process Step Design Use <i>Information resource</i> <ul style="list-style-type: none"> ✓ Metadata database ✓ Units register ✓ Sampling frame Consult <i>Population</i> <p>▶ Output:</p> <ul style="list-style-type: none"> Update <i>Information resource</i> (sampling frame) Create/Update <i>Data set</i> (sample) Create <i>Process Step Execution Record</i> Create <i>Process Metric</i> <p>Supporting Guidelines/Documentation</p> <ul style="list-style-type: none"> ✓ User manual for data marts application ✓ Time table for the given phase of the GSBPM ✓ Instruction for processing 	<p>Transformed outputs</p> <ul style="list-style-type: none"> ✓ <i>Data set: Sample</i> ✓ <i>Information resource:</i> <ul style="list-style-type: none"> • Metadata database • Sampling frame ✓ <i>Process Step Execution Record</i> <p>Process Metric:</p> <ul style="list-style-type: none"> ✓ Sample size ✓ Sampling error <p>Feedback</p>

Phase 8 – Archive

8.1 – Define archive rules

“This sub-process is where the archiving rules for the statistical data and metadata resulting from a statistical business process are determined. The requirement to archive intermediate outputs such as the sample file, the raw data from the collect phase, and the results of the various stages of the process and analyse phases should also be considered. The archive rules for a specific statistical business process may be fully or partly dependent on the more general archiving policy of the statistical organization, or, for national organizations, on standards applied across the government sector. The rules should include consideration of the medium and location of the archive, as well as the requirement for keeping duplicate copies. They should also consider the conditions (if any) under which data and metadata should be disposed of. (Note – this subprocess is logically strongly linked to Phase 2 – Design, at least for the first iteration of a statistical business process).”

Process input	Process step	Process output
Transformable input: <ul style="list-style-type: none">▶ <i>Statistical Program Design</i> from Phase 2 Parameter Input: <ul style="list-style-type: none">▶ n/a Process support: <ul style="list-style-type: none">▶ n/a	(Create/Analyse/Consult/Use/Update/Delete) Process control Business functions <ul style="list-style-type: none">▶ Create/Update Archive <i>Rules</i>▶ Update <i>Statistical Program Design</i> Supporting Guidelines/Documentation <ul style="list-style-type: none">▶ Organization Archiving Policy▶ National Archiving Policy	Transformed outputs: <ul style="list-style-type: none">▶ <i>Statistical Program Design</i> (updated)<ul style="list-style-type: none">✓ Archive <i>Rules</i> Process Metric: <ul style="list-style-type: none">▶ n/a Feedback

The complete work can be consulted in Annex II to this text.

VI. Conclusions and Recommendations

25. Dividing the work into three subgroups didn't work very well because people working on the last phases didn't know what was being defined in the first phases (the work of the three subgroups was in parallel).
26. Working on a 'sprint' basis, with people working physically together, discussing problems, doubts, terminology and concepts seems to be a more productive way to reach a harmonized and homogeneous conclusion.
27. The first difficulty found by the team was working with all the IOs, with different levels of detail; The GSIM IOs should be presented in a more logical format, eventually classified into different levels of detail.
 - a. Examples of these are:
 - Statistical Program, Statistical Program Cycle, Statistical Activity, Acquisition Activity, Production Activity, Dissemination Activity;
28. Some of the IOs have also a structuring role.
 - a. Examples of these are:
 - Process Step, Process Input, Transformable Input, Parameter Input; Process Support input; Process Output; Tranformed Output; Process Metric; Process Control;
29. Identify the higher level information objects in GSIM, and show how the information objects are related;
30. Another difficulty was mapping the terminology statisticians usually used in their offices with the GSIM terminology; This means that, although the definitions of the IOs are in the Glossary, the understanding about them wasn't the same for everyone. These definitions should be improved.

VII. Future Work

31. Taking into account the objectives of this work, the time to do the work and the difficulties experienced, further complementary work might be needed.
32. Explore the relationship between GSIM and GSBPM and establish the flow of the information objects across the sub-process and the role(s) of each information object in each sub-process;

This has been done at a general level (25 main objects).
Is there any value to doing it in more detail (eg for 25 objects<x<=150 objects)?
Maybe at a more detailed level the flows cease to be so generic.
33. Make this relationship understandable to the different populations that deal with these models (managers, specialists), mainly for managers in this phase;
 - ✓ Understand if the current annexes are actually useful to managers and if there is more work that would need to be done on content and/or presentation to make them more useful.

- ✓ Understand also to what extent what has been produced is understandable and useful for specialists and identify what more should be done to meet the needs of specialists (eg in more detail).
34. Test whether the work done on the GSIM model until now fits the purpose for which it was conceived;
35. Regarding GSIM some recommendations are:
- ✓ Analyse all the GSIM IOs and establish which of them are *structuring* IOs, *metadata* IOs or *data* IOs, for example; other different types of IOs can be identified;
 - ✓ Organize information objects in levels of detail;
 - ✓ Being *Statistical Program Design* a high level information object and very relevant to the statistical business process, establish its detailed composition and structure using the GSIM information objects; with this work done, it will be easier to establish which IOs are used in each phase of GSBPM; Identify other cases similar to *Statistical Program Design*, that are high level detail their composition (*Business Case* for example);
 - ✓ Revise the GSIM glossary after the information objects are structured in order to improve and clarify the definitions;
- We present some examples of updates to the definitions included in the GSIM glossary:
- *Data Resource*: an organized collection of data and metadata designed to meet a user need.
 - *Survey Population* and *Target Population* – the difference is not clear by the definition. These definitions should be clarified;
 - *Business Service* – the definition should be improved. It should be clear if software applications are/aren't Business Services.
36. These recommendations could be submitted as Open Issues to the GSIM V1.0 Discussion Forum so they can be reviewed for shaping the next version of GSIM.
37. Evaluate if the GSBPM model needs some refinements;
38. Regarding GSBPM there is one recommendation:
- ✓ Clarify which are the information objects in the business processes (phases 4 to 8) that can contribute to Phase 9 – Evaluation;
39. Contribute to transforming these two models into true standards;
40. Contribute to harmonising terminology across statistical organizations;

These two goals are a consequence of the previous work done. As far as the four first goals are reached the last two goals will be reached through this work as well.

ANNEX I – Metadata flows across the GSBPM (Table 3)

ANNEX II – Metadata Flows in each subprocess of the GSBPM (Table 1)