WP 10
17 February 2010

ENGLISH ONLY

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 (METIS)
(Geneva, Switzerland, 10-12 March 2010)

APPLYING THE GSBPM WITHIN AN NSI: EXPERIENCES AND EXAMPLES FROM AUSTRALIA

Paper prepared by Alistair Hamilton, Australian Bureau of Statistics

I. INTRODUCTION

1. In the past twelve months the ABS (Australian Bureau of Statistics), as Australia's primary NSI (National Statistical Institution), has applied the GSBPM (Generic Statistical Business Process Model) for a range of practical purposes.

2. This paper summarises those experiences. It assumes knowledge and/or reference to the version of the GSBPM that is current at the time of writing (V4.0) [1].

3. In understanding examples of how the ABS has approached, and utilised, the GSBPM to date it is important to first understand the broader perspective of the ABS on the GSBPM and the primary benefits we see its use can deliver. As described in Section II, utilisation of the GSBPM to link metadata management with the statistical business process (the purpose for which development of the GSBPM was originally commissioned by METIS) is seen as just one application of a broader and vitally important “architectural” role for the GSBPM as a reference model.

4. Section III reviews modelling of the statistical business process within the ABS prior to the development and promulgation of the GSBPM. Past experience in this space

- helped the ABS appreciate the design and utility of the GSBPM, but
- led to “sunk investments” in previous “local” (not formally “whole of ABS”) reference models

5. Section IV reviews how the ABS engaged with the development of the GSBPM from V1.0 to the current V4.0. The approach to engaging with the development of the GSBPM is seen as having facilitated acceptance and utilisation of the GSBPM within the ABS.

6. While, as explained in Section II, the ABS views GSBPM's most valuable application relating to a more general, shared, “reference architecture”, Section V explores the GSBPM and its relevance in the context of ABS Enterprise Architecture. The section provides the connection between the historical trajectory traced in Sections III and IV and the specific practical applications of the GSBPM by the ABS to date which are described in Section VI.

7. Section VII outlines some learnings and other outcomes from the initial practical applications and Section

VIII outlines a number of unresolved issues. Section IX identifies some conclusions from an ABS perspective which may be of interest to other NSIs who are seeking to apply the GSBPM.

II. SUMMARISING THE ABS PERSPECTIVE ON THE GSBPM

8. From V1.0 onwards [1], the GSBPM has been recognised as likely to add value as a reference model in the business domain of ABS Enterprise Architecture. The stronger and higher priority focus for the ABS, however, has been the value the GSBPM can contribute as the cornerstone for a more generic "reference architecture" which can be utilised to facilitate and enhance communication and understanding, and ultimately sharing and collaboration, across agencies that are in the business of producing statistics.

9. The role of the GSBPM in such a reference architecture is seen as pivotal. All other aspects of the "reference architecture" (e.g. in terms of applications, data and technologies) ultimately need to be related back to the statistical business purposes and processes they are intended to support. If there is no common reference framework and terminology in regard to the statistical business process then

- elaboration of other elements of the "reference architecture" lacks a defined connection to the actual business of producers of statistics
- proposals to share existing, and collaborate in the development of new, statistical capabilities (e.g., but not only, ICT applications) are less likely to be generated, approved and to result in the desired outcomes because of the difficulty of communicating, understanding and addressing (either as common requirements or via local implementation planning) the business contexts and requirements of each of the possible partners

10. From the ABS perspective, therefore, while the GSBPM itself is generic it is seen as a vital enabler of practical, purposeful and collaborative engagement

- on the international level
  - primarily with other NSIs and international organisations
- on the national and sub-national level
  - primarily with other producers of statistics within Australia's National Statistical Service (NSS) which the ABS is legislatively responsible for leading

11. This had led the ABS to prioritise establishing a reference model which is

- fit for purpose, e.g.
  - able to be understood, and referenced appropriately by, by practitioners within NSIs
  - sufficiently flexible (e.g. avoids being overly rigid and prescriptive)
- commonly agreed and adopted

rather than spending additional time and effort on design and negotiation in pursuit of a "conceptually perfect" model.

12. Given our interest in "common reference architecture" as an enabler across NSIs, and belief in the pivotal role the GSBPM can serve in this regard, the ABS has an intense interest in the activities and outputs of the ESSNET on Common Reference Architecture (CORA), initiated in 2009, whose objective is "to define a common architecture, starting from the GSBPM" [2]. The ABS is aware of, and supports, similar initiatives facilitated by the UNECE which are co-ordinating their work with the ESSNET and which will sustain, and build on, some of the outputs of the ESSNET once the latter initiative is completed later this year [3].

13. The application of the GSBPM by the CORA ESSNET, and related international initiatives, is seen as evidence that the perspectives on, and aspirations for, the GSBPM that the ABS has held from the outset are shared by others.

III. MODELLING OF THE STATISTICAL BUSINESS PROCESSES WITHIN THE ABS PRIOR TO THE GSBPM

14. A key impetus for development of the GSBPM was that when NSIs wanted to discuss metadata in the context of their statistical business processes many agencies had “similar but different” descriptions of their statistical business process and it was not easy to relate the different descriptions to each other. GSBPM would provide a common “reference model” regardless of whether agencies chose to adopt it internally for their own purposes or simply to “map” their internal descriptions to the GSBPM when discussing metadata in the context of the statistical business process with other agencies.

15. The ABS was one of the agencies in this position. In fact, even within the ABS there existed “similar but different” descriptions of the statistical business process. The two most prominent “local” reference models can be (but were not at the time) characterised as the BSIP Reference Model and the ISHS Reference Model. Both of these arose from, and then made crucial contributions to, major programs of business process re-engineering (related to “business statistics” and “household surveys” respectively) undertaken by the ABS during the past decade. Both are described further below.

16. The existence and application of these “local reference models” influenced the ABS approach to application of the GSBPM. In net terms, positive experiences with these “local” reference models facilitated the ABS approach to the GSBPM. The local models had “sunk investments” associated with them, however, that needed to be addressed when considering the future role of the GSBPM within the ABS.

17. The most prominent “local model” was the “BSIP High Level Framework” which served as a cornerstone for Business Statistics Innovation Program (BSIP) initiated within the ABS at the start of this century. This framework was colloquially referred to as “The Caterpillar”.

18. Annex 1 provides more information about the design of the “The Caterpillar”, its application within the ABS and comparisons with the scope and purposes of the GSBPM.

19. A key point of difference is that BSIP and its “Caterpillar” had a strong focus on the ongoing relevance and sustainability of the “statistical business enterprise” overall, in addition to a focus on the statistical business process that was followed by individual statistical activities which produced specific statistical outputs.

20. Documentation of V4.0 states that the “GSBPM is intended to apply to all activities undertaken by producers of official statistics...which result in data outputs” [1]. This means it does not focus on activities whose aim is to ensure the overall relevance and sustainability of the statistical business enterprise (e.g. by providing capabilities that all statistical production processes can harness) but do not directly result in data outputs in their own right.

21. As discussed further in Annex 1, while there is value in common reference frameworks that relate to the “statistical business enterprise” as a whole (which are known to already exist to some degree but possibly not in as contemporary and succinct form as the GSBPM) it is seen as a positive rather than a negative that the GSBPM itself does not attempt to fulfil this extra role.

22. Nevertheless, for some in the ABS it has taken time to get use to this deliberate difference in scope between the GSBPM and the Caterpillar.

23. Synergistically, alongside BSIP, there was a more general focus on describing and managing Enterprise Architecture (EA) within the ABS. This led to The Caterpillar often being used as a reference model when describing business architecture well beyond BSIP.

24. This use was premised on the belief that, similarly to the GSBPM, The Caterpillar was broadly applicable to all statistical activities which produce data outputs, not only to business surveys within the scope of BSIP and to the integration of alternative administrative sources in that context.

25. Since that time, however:

- agreement could not be reached that the Caterpillar provides a good common starting point for modelling compilation activities (e.g. national accounts, price indices)
- the Integrated Systems for Household Surveys (ISHS) initiative chose a “similar but different” reference model (which used the Caterpillar as a starting point, but then diverged)
- a number of other key statistical activities (e.g. Population Census, various demographic and social statistical activities utilising data from administrative registers) did not see strong value in describing their statistical business processes using The Caterpillar
- corporate “staff induction” training in regard to the business of the ABS continued to utilise an earlier cyclical generic model of the business process

26. In terms of applicability, there has been no questioning within the ABS that the GSBPM is applicable as a reference model to both business and household surveys. (The debate on the extent to which this means GSBPM should now within the ABS “supplant” the separate reference models developed for BSIP and ISHS is described the following section)

27. While it has been broadly agreed within the ABS that, by

- omitting some “survey oriented” sub-processes, and
- applying a broad interpretation to others

the GSBPM could be used more generally as a reference model for non survey based statistical activities – e.g. price index compilation, accounts compilation – whether this would prove useful in practice has been left as an open question. This hesitation is partly a legacy of ABS experiences with the Caterpillar, and is open to being addressed through a successful overseas or local practical application of the GSBPM to such activities.

28. Annex 2 outlines the value realised by the ABS from application of these local reference models. It highlights that where Section VII of the GSBPM 4.0 documentation describes “Other Uses of the GSBPM”, most (but not all) of the uses listed were to at least some degree validated within the ABS using the BSIP and/or ISHS reference models.

29. The uses that rely on international and/or corporate application of the GSBPM are, unsurprisingly, among the uses that were least well demonstrated by application of the BSIP and ISHS reference models.

30. In regard to corporate applications, while there was not a formally agreed and applied translation between the BSIP and ISHS reference models, their common heritage meant that for some purposes the differences between them were not significant. For example, at a corporate level when considering in general terms costs, or management of quality, it was possible to group information and discussions around elements from the BSIP reference model even though some organisational units used a different model, or none at all, for their internal purposes.

31. In summary, each reference model was invaluable to business process re-engineering program that gave rise to it, including to the cluster of statistical activities associated with each program. Without being formalised across the organisation as a whole, the BSIP reference model added significant value more generally as well.

IV. ABS ENGAGEMENT WITH THE DEVELOPMENT OF THE GSBPM

32. Annex 3 summarises ABS engagement with the development of the GSBPM from V1.0 to V4.0

33. A factor that appears to have facilitated acceptance and utilisation of the GSBPM within the ABS is a progressive broadening of the audience engaged to provide feedback on the process model as it evolved over time. Any attempt to start by engaging the organisation as a whole in regard to V1.0 may have been premature, generating more confusion than consensus. The fact there was broader engagement prior to V4.0, however, avoided other stakeholders within the ABS feeling they were being introduced to a “final” model that had been developed for use without recognition of their perspectives and needs.

34. As described in Annex 3, initial interest was from the ABS divisions with responsibility for statistical methods (including the management of data and metadata) and for ICT [4]. Both of these recognised the value that could arise from further development and promotion of the GSBPM, with the aim of facilitating its future application for various purposes both internationally and within Australia. Both subsequently acted as proponents of the GSBPM within the ABS.

35. When introduced to the GSBPM business representatives were initially more reserved in regard to how it might be applied. They were willing to "map" to GSBPM for reference purposes outside the ABS but were sceptical about its application as a reference model within the ABS. They suggested that, if and where the need arose to look across the ABS as a whole, other models used within the ABS could be mapped to the BSIP Caterpillar for reference purposes.

36. Many of the reservations seem to have arisen for historical reasons.

37. A significant concern was that over several years staff in large parts of the organisation had become familiar with categories and terms associated with the BSIP or ISHS reference models.

38. As described further in Annex 3, there were also a range of more tangible "sunk investments" related to the BSIP and ISHS reference models (e.g. in terms of the structure and content of documents, diagrams and, in a number of cases, user interfaces to applications and repositories delivered by the programs).

39. While there is agreement on the need to recognise and address these sunk investments as part of any decision to apply the GSBPM within the ABS, the details of implementation guidelines in this regard are yet to be finalised.

40. In response to a more recent UNECE invitation for the head of the ABS to provide for feedback on the GSBPM a larger and more diverse group (in terms of the range of levels and roles) within the organisation was consulted.

41. The feedback supported

- the relevance and value of the GSBPM as a reference model when communicating across agencies
- the current level of detail provided within the GSBPM

42. The major unresolved question identified in that feedback, however, was how the ABS (and others) can continue to engage with the ongoing evolution, extension and international application of the GSBPM. This is discussed further in Section VIII.

V. POSITIONING THE GSBPM IN THE CONTEXT OF ABS ENTERPRISE ARCHITECTURE (EA)

43. Various EA frameworks exist. The ABS now applies TOGAF 9 [5].

44. Annex 4 provides further information about EA within the ABS and the position of the GSBPM in regard to it.

45. There is a clear connection between ABS interest in GSBPM's cornerstone contribution to more generic "common reference architecture" (as outlined in Section II) and its status within ABS EA. While a level of interoperability supported by "common reference architecture" should be possible even where none of the reference models and standards have been "internalised" by any of the enterprises concerned (leaving each enterprise with the task of mapping "from first principles" its internal approaches to the external reference approach and vice versa), interoperability is expected to be more efficient and effective if the reference models and standards are also harnessed within the enterprise to the extent they are fit for purpose. The GSBPM is viewed as fit for many purposes as a reference model in the business domain within ABS EA.

46. During discussions in the last quarter of 2008 and the first quarter of 2009, representatives of statistical business areas informally concurred with the two “proponent” divisions (ICT and Statistical Methods) that the GSBPM is useful as a reference model when conversing beyond the boundaries of the ABS in regard to the statistical business process. They did not concur that it should be the preferred reference model for use within the ABS.

47. Following the METIS Workshop intended to “finalise” the GSBPM in March 2009 [6], and the release of V4.0 the following month, there was an upswing in ABS interest in the GSBPM. This was primarily driven by the GSBPM's new level of maturity as an internationally agreed (“de facto standard”) reference model for the statistical business process. (The progress in content between V2.0 and V4.0 was perceived as less of a driver for wider adoption than the progress in status.)

48. By this time the ABS framework for EA was undergoing fundamental review, meaning that, contrary to previous expectations, the change in external status did not immediately translate to a formal upgrading of the status of the GSBPM as a reference model within the ABS. The decision to adopt TOGAF has been one outcome of this review, together with changes in approach that more fully engage the whole organisation in defining, owning and applying EA.

49. It is expected that as part of implementing this new approach to EA, an early step will be formal corporate recognition of the GSBPM as a preferred reference model for the statistical business process.

50. In the meantime, initial ABS applications of the GSBPM have been associated with the two “proponent” divisions.

VI. PRACTICAL APPLICATIONS OF THE GSBPM BY THE ABS

51. The first applications of the GSBPM (in diagrams and documents rather than in processes and systems) utilised an earlier version (V2.0) of the model.

- As leader of the National Statistical Service (NSS) within Australia, the ABS makes many references to the statistical business process when sharing ideas, good practice and directions with other producers and users of statistics, and also when providing more detailed support or planning collaboration. Selected NSS diagrams and documents began to incorporate the GSBPM.
    - An added advantage is that the GSBPM is seen, by the ABS and by others, to provide a neutral and external framework (a de-facto international standard) rather than an “ABS-centric” one
    - This was reinforced when another major producer of statistics (the Australian Institute of Health and Welfare) independently decided to apply the GSBPM when modelling their statistical business processes
- In the second half of 2008 an initial proposal was developed for a new approach to statistical information management throughout the ABS statistical business process, and beyond the boundaries of the ABS, harnessing standards such as SDMX and DDI. Once again, as the proposal was intended to engage with stakeholders beyond the ABS, the GSBPM was used to represent the statistical business process.

52. These early uses explicitly noted the GSBPM was yet to be finalised.

53. More recent applications by the ICT division include

- In support of a recent major ICT Strategy review a map printer was used to produce a giant representation of the GSBPM. This was then used as one focal point for brainstorming sessions with stakeholders from across the organisation. The GSBPM wall chart soon became covered in post-it notes highlighting issues, opportunities and gaps in regard to ICT support for end to end statistical business processes within the ABS.
- In a related portfolio review of the health of existing applications (both in terms of technology platforms and alignment with business needs) the GSBPM was again used as a generic means of

categorising the business processes supported by the application. This provided a starting point for exploring how well the application fulfilled these roles and how well it interoperated with “upstream” and “downstream” applications. More importantly, when looking across the portfolio it helped highlight phases and sub-processes where there may be

    - an unnecessary proliferation of different applications being applied and maintained currently
    - no existing application that is “in the prime of health” from both technical and business perspectives
- New capital expenditure proposals now refer to the GSBPM as part of documenting their expected fit, and value add, within end to end statistical business processes within the ABS.
- Early consideration is being given to how the phases of the GSBPM will be used as references and an “organising principle” when implementing the new Business Process Management System within the ABS.

54. The GSBPM has also been utilised heavily in early planning and communication associated with the ABS Information Management Transformation Program (IMTP). A key aspect of IMTP is the application of open standards such as DDI and SDMX to address current “disconnects” in the flow of data and application of metadata

- between sub-processes within the ABS
- between the “internal” world of the ABS and the “external” world (e.g. respondents, administrative data sets, end users of statistics)
- within the “external” world of Australia's National Statistical System (supporting flows and applications where the ABS may be neither the content producer or consumer)

55. In regard to the IMTP

- The role of the GSBPM as a reference point has been instrumental when formulating, negotiating and communicating strategic decisions about “where to start”.
- Individual pathfinder projects have been represented in detail in regard to sub-processes within the GSBPM, both in terms of the focal points for the project itself and in terms of critical touch points with other processes (e.g. dependencies, downstream ramifications).
- Work is just beginning on a more detailed map of relationships between different types of statistical metadata objects (classed according to the information model used by DDI) and the statistical business process (classed according to GSBPM).
    - The mapping is complicated by the fact that relationships may have many forms, such as
        - creation of fundamentally new metadata objects
        - reuse of existing metadata objects
        - updating (e.g. elaborating upon) existing metadata objects
        - linking together existing metadata objects in a new way for a new purpose

56. The current ABS focus is on the relationship between DDI and GSBPM because

- this appears not to have received the same attention to date as relationships between SDMX and GSBPM
- DDI is seen as more directly relevant than SDMX to many GSBPM phases within the ABS

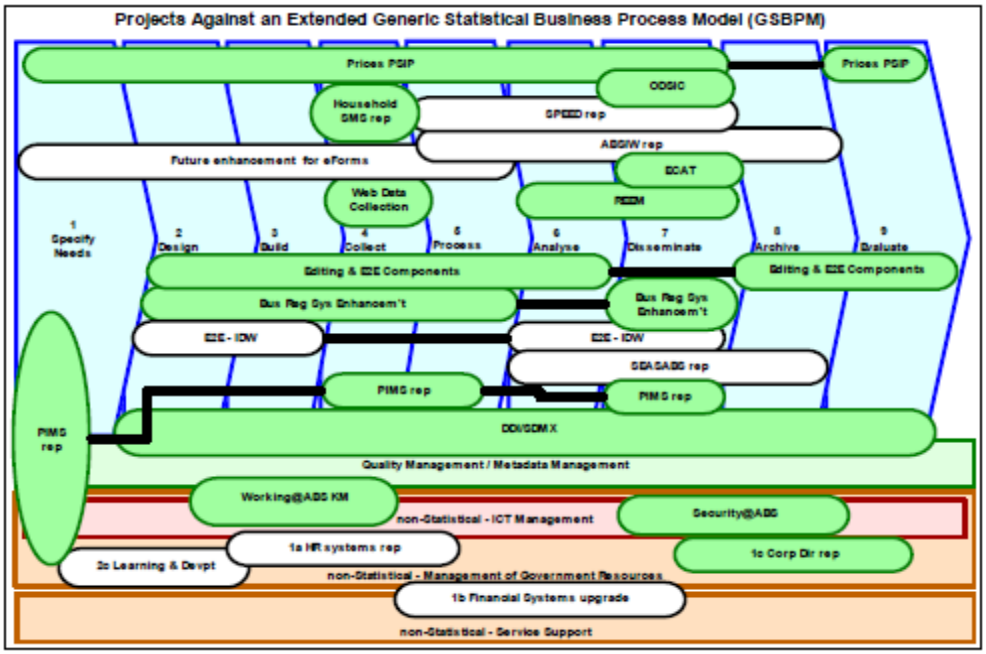
57. Nevertheless, the ABS also has a practical interest in connections between SDMX and the GSBPM.

VII. EARLY LEARNINGS

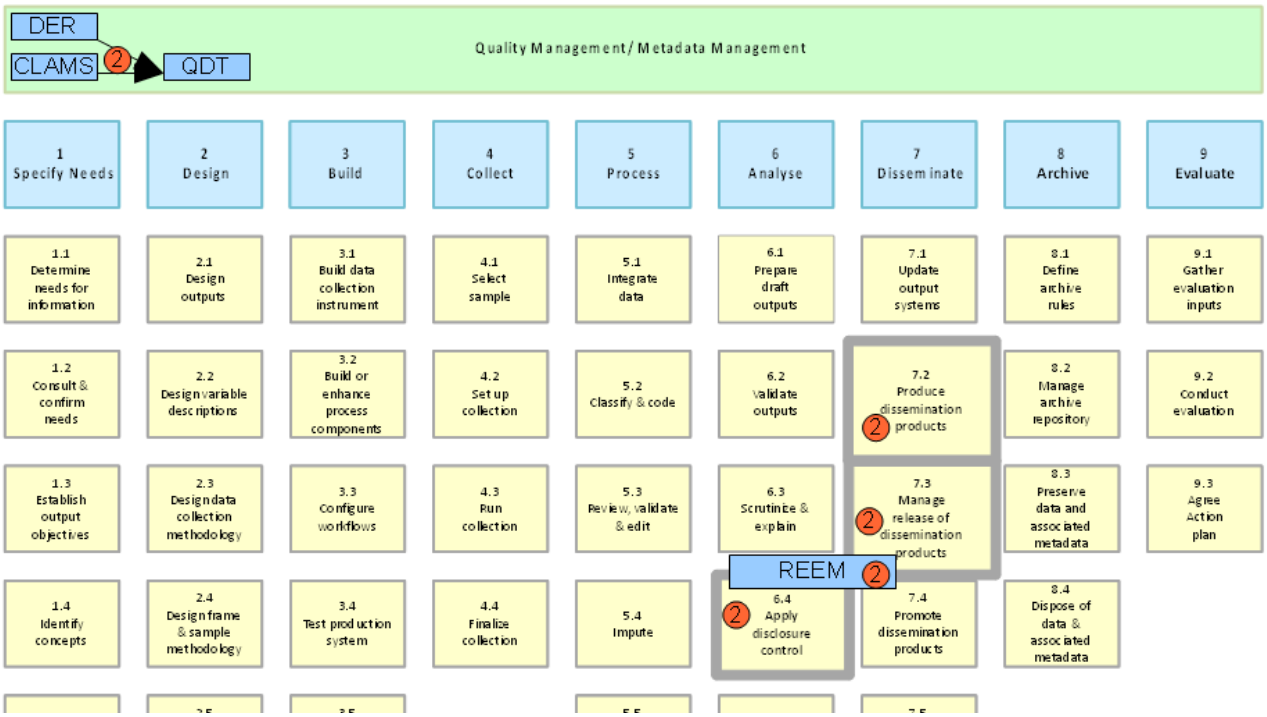
58. Positioning and illustrating specific existing work flows, applications and projects relative to the GSBPM seldom produces a “neat” picture. A cluster of sub processes within a particular phase may “light up”, but other sub processes from much earlier or later phases may also form part of considerations, without

necessarily “connecting” via activities in contiguous phases.

59. The following diagram is deliberately not legible in terms of individual projects but illustrates the idea, with the heavy black lines indicating connections between “islands” of relevance to various phases.



60. Once again, without going into detail about the specific project, the following illustrates the pattern for a project at the sub process level which has significant implications for four phases and four applications (which are identified by acronym in the blue rectangles)



61. Even though the results are often disjointed, mapping against the GSBPM has proved very valuable. In some cases it highlights that applications

- are being used for business purposes outside the original scope of their design, or
- are not being used in practice for purposes that they were intended to be able to support

62. In the case of workflows and projects, mapping is useful to help

- ensure important components, connections and implications are not overlooked
- ensure manageable and commonly understood and agreed scoping through identifying which aspects will be a primary focus and which will be recognised as "related/relevant but not central"
- picture, in broad terms, how a set of projects or set of workflows relate to each other

63. Although illustration and/or categorisation using the GSBPM is usually very valuable for orientation in regard to workflows, projects, applications etc, and for promoting more detailed questioning and discussion, even high level consideration of their relationship with statistical business processes often benefits from additional explanation, contextualisation and detail beyond that provided by the GSBPM on its own.

64. It has long been recognised, within the ABS and beyond, that in order to undertake more detailed planning, design and collaboration for a particular purpose there is a need to go to a greater level of detail and specificity in process specification ("as is" and "to be") than is provided by sub-processes and their description within the GSBPM.

65. Experience to date has confirmed this. The often disjointed results from mapping to the GSBPM have further suggested that for the purposes of a specific undertaking it may not always be appropriate to use GSBPM sub-processes and phases as the main organising principle for process specifications. For example, the primary common "working representation" of the business processes in scope, and how they fit together, to be used by the project team may not have the nine phases of the GSBPM, or a subset of them, as its backdrop.

66. Nevertheless

- terminology associated with the GSBPM should be reused wherever relevant, and should not be used in a manner that is inconsistent with its usage in the GSBPM
- the project team should be clear on the relationship between each component of their "working representation" of in scope business processes and the GSBPM
- the team should still provide a "standardised" summary view in terms of the GSBPM

67. Early experiences have tended to allay concerns expressed in some quarters previously that the GSBPM would prove too general and/or too inflexible to be of practical value when applied within the ABS.

68. As further highlighted in the next section, however, these early applications have also emphasised the need to consider carefully where, and how, the GSBPM adds value as a primary frame of reference. There are other cases where it could be applied, but would not be the best choice.

VIII. UNRESOLVED ISSUES

69. ABS experience to date in regard to application of the GSBPM has been very positive. Nevertheless, the ABS recognises a number of issues that will need to be addressed either by us or by others in order to further promote the application of the GSBPM and to obtain maximum value from its application.

70. It is considered worth summarising the issues to be resolved within the ABS, in addition to those where primary carriage would not be with the ABS, as other NSIs may face similar implementation issues and sharing of ideas and experiences my be worthwhile.

Promoting understanding and application of the GSBPM within the ABS

71. As described in the previous sections, the GSBPM has already been applied for various purposes by two "proponent" divisions within the ABS. As a result of applying the GSBPM for these purposes many more people, including managers and staff in statistical business divisions, have encountered graphical representations based on the GSBPM and the names of phases and sub-processes.

72. This does not imply, however, there is now a common understanding across the ABS of the GSBPM and its application. Many staff would not know

- when they should utilise the GSBPM as a frame of reference (compared with, e.g., the "local" BSIP and ISHS reference models)
- how to utilise the GSBPM, e.g.
  - they may have seen the phase labels but not the sub-processes and their descriptions and so may not have a clear sense of scope when applying phase labels
  - they may not know where to locate the V4.0 documentation and/or how to resolve any remaining questions after reading it

73. Formalising the status of the GSBPM with respect to ABS Business Architecture, as described in Section V, is seen as a key trigger in that regard.

74. Corporate communication in the meantime would be either be based on the perspective of the two proponent divisions or have as a prerequisite the same level of engagement across other divisions which is required to formalise the status of the GSBPM within ABS Business Architecture.

75. It is not expected building this understanding would entail "heavy duty" communication and training. The keys would be

- a summary of the GSBPM itself, including the value of it to the ABS and others, and a link to the full V4.0 documentation
- clear and concise guidelines (e.g. a couple of pages) on applying the GSBPM within the ABS
- an ABS point of contact for further information, assistance and feedback

76. A clear point for resolution in the EA discussions, and communication subsequent to that, will be the status of the GSBPM relative to the BSIP and ISHS reference models. It has already been agreed existing process documentation will not be rewritten for the sole purpose of referring to the GSBPM. Questions yet to be finalised include

- the extent to which business areas associated with BSIP and ISHS should start orienting new content (e.g. documentation, diagrams) toward the GSBPM, thereby creating a "break" with their past material
- the extent to which existing content should be re-categorised/remapped to align with the GSBPM even where the content itself is not rewritten
- whether minor "refreshes" of content for other reasons will necessarily trigger a need to update that content to refer to the GSBPM

77. Formalising, and making readily available, the existing mappings between the GSBPM and the local reference models will also be a key aspect.

GSBPM and the DDI Life Cycle Model

78. Since V2.0 the GSBPM has included references to the DDI Life Cycle Model (LCM) and discussion of mapping between the two.

79. A decision taken in October 2009 to implement DDI V3 has focused additional attention on this relationship within the ABS.

80. As explained further in Annex 5, the ABS sees applications for both models, and agrees they should remain distinct, but would like to see as much interoperability as possible between the two.

81. Addressing unresolved issues in this case probably entails the ABS and/or other NSIs and/or the UNECE having discussions with the DDI Alliance.

82. Firstly, as increasing numbers of NSIs are showing interest in DDI V3, it would seem positive for the Alliance to consider the GSBPM more closely (although not as a replacement for the LCM, nor as a reference model that is likely to be relevant to all members of the DDI community).

83. As described in Section VI, the ABS sees value in, and has commenced producing, an output that seeks to relate objects defined within the DDI information model to typical usage within statistical business processes (with reference to the GSBPM). Advice from the DDI Alliance that the output appeared reasonable would help validate, at least in a general sense, the ABS interpretation of the DDI information model and its application within our business environment as an NSI. This output, having been broadly quality assured by the Alliance, might also then prove useful for other NSIs who are looking to engage with the DDI standard.

84. Secondly, having looked further at DDI in the context of the GSBPM, it might be appropriate to confirm from the perspective of the DDI Alliance that they are comfortable with the current mapping between the DDI LCM and the GSBPM. This might include considering whether there should be more detailed definition of the DDI LCM itself and the terminology used within it. (While overall DDI is a much broader and more formal standard than the GSBPM, the DDI LCM does not appear to be as well defined as the GSBPM.)

85. The current mapping between the GSBPM and the DDI LCM provided in the documentation of V4.0, for example, suggests “Disseminate” is synonymous with “Data Distribution” and “Archive” is synonymous with “Data Archiving”. This may be worth revisiting because, for example, it has been clear though recent discussions with the Alliance that how the ABS traditionally conceives “Archive”, and how the DDI Alliance conceives “Data Archiving”, are significantly different in scope and nature. (The ABS addresses “Data Archiving” in the sense used in DDI, but we don't associate this with the verb “Archive”.)

86. If, after revisiting, it is confirmed that the terms used in the two models are synonyms then it would be good to work toward common terminology. If there are meaningful differences then these should be captured as part of the documentation of the relationship.

87. The mutually validated relationship between the GSBPM and DDI LCM might then be set out in a separate document to either the GSBPM or the DDI LCM themselves. The documentation of the mutually agreed mapping would then be referenced from the GSBPM and possibly from the DDI LCM.

88. If there was any practical value in doing so, in addition to textural descriptions there could also be machine actionable DDI representing both models as schemes and also a machine actionable representation of the relationship between the two.

Establishing an informal community around the GSBPM

89. As the GSBPM becomes increasingly used for reference purposes within the ABS, within other NSIs and across NSIs the importance and value of establishing a “community of practice” around it grows.

90. The ABS strongly supports community driven open standards such as SDMX and DDI. Such standards are open not only in the sense that the outputs (various versions of the standard, supporting material and sometimes supporting infrastructure) are open and accessible but aspects beyond formal outputs such as

- proposals for evolution and extension
- ideas, experiences and questions fed back from practitioners applying the current standard

are also to some extent accessible to an open community of interest around the standard. Such an approach is seen as beneficial both

- to the standard itself, e.g.
  - evolution is better informed by the interests and needs of the broader community
  - there are higher levels of adoption and support
  - there is an ability to draw on the resources of the community to test and validate candidate versions and proposed extensions
- to the community which utilises the standard, e.g.

- practitioners are not limited to simply “consuming” the current outputs but have a real opportunity to influence what is delivered in future
- members benefit from understanding experiences and plans related to the standard from other community members whose context may be similar to theirs
  - this is valuable in its own right, but can also lead further to further sharing of information and even active collaboration in regard to areas of common interest
- while the standard will have boundaries (e.g. where it “hands off” to other standards rather than covering the same territory as them) the interests of many in the community can extend, and work across, those boundaries
  - members of the community gain insight to, and can collaborate in regard to, other standards that may be relevant to them and how to best harness them in combination

91. While the GSBPM is not of the same scope and nature as the information models that underpin standards such as SDMX and DDI, it is seen as an important and useful reference model that was developed by a community centred on “official statistics”, including NSIs, to serve that community.

92. The ABS hopes, and expects, many similar benefits to those dot pointed might be realised were an informal community of practice to be supported around the GSBPM.

93. While Part C of the Common Metadata Framework (CMF), which provided the original impetus for the development of the GSBPM, continues to make use of it, the GSBPM now has its own identity within the METIS wiki, independent of the CMF [7]. Similarly to the way Part C of the CMF references the GSBPM for its purposes, a number of other international and national uses do the equivalent already.

94. This separation can be seen as fortuitous when considering interest in the GSBPM from a wider community.

95. Given the nature of many of the uses set out in Section VII of the V4.0 documentation, many of those with an active interest in the GSBPM may not have a particular interest in the CMF.

96. It might be asked whether a community related to the GSBPM should be hosted within METIS wiki at all.

97. For example, as mentioned in Section II, one of the most exciting uses of the GSBPM to date from an ABS perspective is its adoption as the prime reference model for the ESS Common Reference Architecture (CORA) Project. The CORA questionnaire has generated a lot of information aligned with the GSBPM. Furthermore it is anticipated that its use within the questionnaire will have increased general awareness and interest in regard to the GSBPM.

98. The CORA ESSNET has a finite life, however, and its outputs will be maintained by another UNECE convened group. As a number of UNECE convened groups, including METIS, will have an ongoing interest in the GSBPM it is not seen as a vital issue whether it continues to be hosted within METIS, within a different group or in its own right.

99. What is proposed is that beyond the single page that now hosts the current version and past versions of the GSBPM, an opportunity be created for an informal community of interest. Supporting the community within the current wiki might be one option (e.g. a similar approach to that currently taken with the metadata case studies [7] but less formal?). This would not be the only option.

100. Section I of the V4.0 documentation states

> *it is also expected that future updates may be necessary in the coming years, either to reflect experiences from implementing the model in practice, or due to the evolution of the nature of statistical production.*

101. The commitment to evolution based on practical experience and on a changing context is strongly supported.

102. The community should be a primary source of information and ideas in this regard.

103. It is much more effective to record issues and ideas as they arise rather than trying to gather them months or years later through periodic requests for input toward a new version. Also, as the community of interest around the GSBPM grows and diversifies it becomes increasingly problematic to target periodic requests for input toward a new version. A request for input and/or a draft of a proposed new version could be placed on the current web page but, because that page is static for months at a time – and does not link to more active pages in regard to the GSBPM, many who have an interest in the GSBPM may have downloaded the PDF and not refer back to that page regularly.

104. Pending updates to the GSBPM, a community can also help limit divergence. For example, if two or more agencies are finding it difficult to categorise a particular detailed business process within the high level framework they may be able to agree on a common interim approach and on a proposed update to the next version of the GSBPM. While this will not eliminate divergence in approaches it may help reduce it. It will also give people a sense of shared experience around understanding and implementing the GSBPM, making it seem less "impersonal" as a model.

105. Even more importantly, however, supporting a community that was able to share ideas, and information about activities and experiences, related to the GSBPM would be an organic “bottom up” means of helping realise the potential of the GSBPM as a key underpinning of a common reference architecture, and the enhanced opportunities for sharing and collaboration associated with that. This would complement the structured “top down” approach taken by the CORA ESSNET and similar initiatives.

106. Also, for example, when the first practitioner from another NSI shares their techniques and learnings from, e.g., applying the GSBPM as a reference model to a statistical business processes involving administrative data sources, national accounts compilation or similar that is likely to further broaden interest and confidence within the ABS.

107. It is recognised that establishing and supporting such a community takes effort, potentially significantly increasing the already heavy workload of the UNECE secretariat. Where it is feasible to share aspects of the additional administrative load beyond the UNECE, the ABS would be happy to contribute to the effort.

108. In addition, to ensure some content is available from the outset, the ABS would be happy to contribute one or more informal case studies related to application of the GSBPM to date.

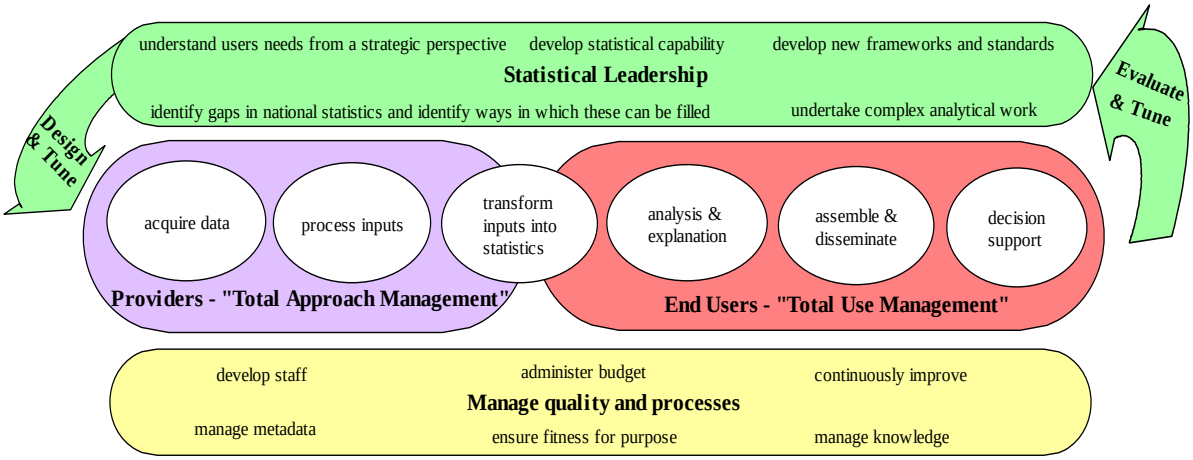
IX. CONCLUSIONS

109. Key points from the ABS experience include

- a belief in the value of the GSBPM as a preferred reference model for application within an NSI, in addition to its broader role as a reference model that facilitates communication, sharing and collaboration across NSIs
    - for different purposes, the ABS also sees ongoing value in the DDI Life Cycle Model
    - other agencies may well conclude the same in regard to other locally developed and/or internationally sourced reference models that are in use currently
- where there are reference models already in use within an NSI, any proposal to harness the GSBPM needs to be thought through, and communicated, carefully and with full recognition of these pre-existing reference models and any “sunk investments” related to use of them
- a recommendation that, based on its various uses, the GSBPM be regarded, and be approached, as a resource of interest to the organisation as a whole rather than being just the province of the “metadata management people” and/or “the ICT people”
- when introducing GSBPM to a wider audience it is important to emphasise its nature as a reference model rather than it being intended as a prescriptive framework for assembling the actual business processes used within an NSI

    - the documentation already makes this clear, but
        - some staff are not familiar with the concept and application of reference models
        - some staff are keen on prescriptive frameworks and will seek to interpret and apply any model accordingly
- likewise the understanding needs to be clear that while application of the GSBPM should facilitate communication, and it is hoped sharing and collaboration, in regard to business processes it is not anticipated that reference to the GSBPM, in itself, provides all the information about processes required to make this happen
    - once again, the documentation already makes this clear but misapprehensions about what the GSBPM sets out to achieve have been common in discussions within the ABS

Annex 1 : BSIP High Level Framework (aka "The Caterpillar") vs the GSBPM



The Caterpillar provided a vital underpinning, including as a common point of reference and as an "organising principle", for Business Statistics Innovation Program (BSIP) which was initiated within the ABS at the start of this century.

While not universally recognised and adopted across the ABS, "The Caterpillar" was more fully developed, more extensively utilised in practice and more widely shared across the ABS than various previous models of the statistical business process that had been created for various purposes since time immemorial.

The Caterpillar emphasises the "cyclical" nature of many statistical activities and, in keeping with the themes of BSIP, the need to continuously tune not only the "mechanics" of the process but the degree to which its outputs meet the statistical needs of users in the context of the wider suite of data that is available to them.

A key emphasis for BSIP was not just undertaking existing statistical activities more efficiently and effectively but determining which activities (whether existing or new, whether undertaken by the ABS or by others) were most important to users of statistics within Australia and reviewing how best to undertake these activities (e.g. moving from survey to administrative sources in some cases, scaling back in some domains of decreasing relative interest to allow addressing of fundamental unmet needs in other domains).

There was also an emphasis on maintaining capabilities (e.g. in terms of skills, knowledge, finances, technology) to support statistical activities, with the expectation those activities would change over time more rapidly both in broad nature and in detail than had been the case previously.

Both of these areas of emphasis, which could be broadly summarised as relevance and sustainability, are essential to the business of an NSI. While all enterprises, especially in facing the challenge of an ever more interconnected and rapidly changing operating environment, need to pay attention to their ongoing relevance and sustainability, how NSIs approach this is strongly shaped by the fact that our core business output is official statistics rather than something else.

While these two areas of emphasis were reflected in the Caterpillar, they are largely placed outside the scope of the GSBPM which is "intended to apply to all activities undertaken by producers of official statistics...which result in data outputs." Relevance and sustainability tend to be managed "across" the activities which produce data outputs at least as much as they are managed "within" individual data production activities. The simplicity, and possibly some of the general applicability, of the GSBPM would be diminished if it sought to extend into relevance and sustainability in more detail.

This difference, however, initially caused some confusion within the ABS when the GSBPM was compared with the Caterpillar.

When considering metadata in the context of the overall statistical business of an NSI, a critical component is its contribution to those aspects of relevance and sustainability highlighted in the Caterpillar but not highlighted in the GSBPM. In the context of the METIS Common Metadata Framework [7], these aspects can

be seen as addressed primarily via Part A – Statistical Metadata in a Corporate Context rather than via Part C – Metadata and the Statistical Business Process. It is Part C where the GSBPM is utilised as a reference model.

There is a common tenet within the ABS that “The Caterpillar” influenced Statistics New Zealand's subsequent design of their generic Business Process Model (gBPM) which, in turn, provided the primary basis for the GSBPM. Regardless of the extent to which this tenet is founded in fact, the common assertion that GSBPM shares DNA with The Caterpillar appears to have facilitated ABS uptake!

Annex 2 : Value realised from pre-existing statistical business process models

In the course of the past decade the practical value and application of the BSIP reference model (the Caterpillar) and the ISHS reference model has been proven within the ABS.

The BSIP and ISHS programs targeted rationalisation and modernisation of business processes and supporting infrastructure for key “clusters” of statistical activities (related to business statistics and to household surveys respectively). These initiatives entailed a lot of business process mapping of “in scope” statistical activities, both in “as is” and proposed “to be” forms.

In the context of these initiatives, a great deal of value was added specifically by the respective business process reference models. In fact, use of the reference models was so prominent that a graphic of the model was frequently used as a logo, or other form of symbol, for the program as a whole.

For example, the reference models

- were essential to support consistent and meaningful discussion of in the aims and implications of these business process re-engineering programs (and of individual projects within the programs) with managers and staff who were used to thinking in their own terms about their own particular statistical activity

- were essential to ensuring a starting point for comparability, analysis and planning and co-ordination of change across the different existing “in scope” statistical activities when capturing each activity's existing process steps in the form of an “as is” business process map

- helped free the definition and design of new shared capabilities (e.g. methods and supporting infrastructure) from being visualised in the context of a specific existing process

- helped organise information related to the initiatives, their inputs and their outputs,
    - for example, the following information resources were often structured and/or indexed based on headings from the relevant reference model
        - collections of documentation about the current business process steps and the transition plans for individual statistical activities
        - “business process guides” (e.g. recommended practices, lists of useful resources) resulting from the initiatives that were intended to be relevant to any in scope statistical activity
        - documentation set in place for individual statistical activities related to their particular “post transition” set of business process steps

Annex 3 : ABS engagement with the development of the GSBPM from V1.0 to V4.0

A factor that appears to have facilitated acceptance and utilisation of the GSBPM within the ABS is a progressive broadening of the audience engaged to provide feedback on the process model as it evolved over time.

Version 1.0

V1.0 was discussed in some detailed within the ABS divisions with responsibility for statistical methods (including the management of data and metadata) and for ICT.

The timing was fortuitous as it coincided with the ICT division having just decided to embark on a rationalisation across the BSIP reference model (i.e. the Caterpillar), the ISHS reference model and the gBPM from New Zealand to seek to establish a true "whole of ABS" (i.e. enterprise level) reference model.

It was quickly agreed that effort should, instead, be focused on contributing to the international effort (e.g. through providing feedback on V1.0 of the GSBPM) and then seeing if the results could also be applied for "whole of ABS" purposes.

Version 2.0

In addition to the Statistical Methods and ICT divisions, the areas that had been responsible for leading the BSIP, ISHS and Census 2006 business transformation projects were engaged more closely in these consultations within the ABS. These discussions confirmed

- the BSIP and ISHS reference models were readily mappable to the GSBPM
- the GSBPM offered some advantages over either of the "local" reference models
    - from the point of view of the business representatives, however, these were not seen as sufficient to warrant moving from their "local" model to use of the GSBPM

The business representatives were willing to "map" to GSBPM for reference purposes outside the ABS but were sceptical about its application as a reference model within the ABS. They suggested that, if and where the need arose to look across the ABS as a whole, other models used within the ABS could be mapped to the BSIP Caterpillar for reference purposes.

Many of the concerns seem to have arisen for historical reasons.

One set of concerns arose from the fact the two local reference models had come to be identified with the particular business transformation programs that had given rise to them. As described in Annex 2, each reference model had fulfilled a vital role in supporting communication, co-ordination and planning within the associated transformation program.

There was a concern that, by its nature, the GSBPM would be too generic, or otherwise unsuitable, to fulfil a similar purpose in future for any programs as intensive as BSIP and ISHS which might target a specific "stream" of statistical activities.

In this regard, it took time to recognise the GSBPM as a further level of abstraction. Adopting the GSBPM as a primary reference model internationally and/or within the ABS would not include the requirement that any future business process transformation program use the GSBPM directly as its "day to day" reference model for supporting communication, co-ordination and planning across the statistical activities and other stakeholders that were in scope for the program.

If necessary a future business transformation program could define its "to be" business process model with reference to the GSBPM but include more context specific descriptions of the relevant phases and sub processes and go to extra levels of detail below sub processes. Projects and statistical activities in scope for that program could then refer primarily to the program level "reference model", without losing the connection

to the more generic reference model.

Less simply addressed were concerns that over several years staff in large parts of the organisation had become familiar with categories and terms associated with the BSIP or ISHS reference models.

Even more significant in terms of “sunk investments” associated with the BSIP and ISHS reference models were factors such as the following :

- some systems and their user interfaces had been built with reference to one or other of these models,
- a lot of documentation had been assembled and categorised based on these models, and
- even more documentation (including many diagrams not amenable to a simple “search and replace” approach) referred to one or other of these models either by name, graphically or via the choice of terminology

Arising from this round of consultation the ABS provided extensive comments on V2.0.

As described in Section IV, following these internal discussions of V2.0 the ABS also started to make practical use of the GSBPM in contexts where reference to an ABS specific model of the statistical business process was not appropriate.

Version 3.1

V3.1 was only current for three months. It was drafted in response to feedback on V2.0 and designed as input to a METIS workshop, held in March 2009, aimed at “finalising” the GSBPM.

An initial assessment suggested V3.1 introduced no new concerns from an ABS perspective and that it addressed several elements of the feedback provided on 2.0.

Under those circumstances, having undertaken relatively extensive consultation within the ABS on V2.0, it was considered an inappropriate use of a wider group's time to convene a further extensive round of consultations within the ABS on V3.1.

Version 4.0

This was the product of the METIS workshop mentioned above.

In August 2009 UNECE invited feedback on the GSBPM from members of the Bureau of the Conference of European Statisticians, one of whom heads the ABS. In order to ensure more than a repeat of comments offered as a result of previous consultations, a larger and more diverse group (in terms of the range of levels and roles within the organisation) was consulted.

The feedback supported

- the relevance and value of the GSBPM as a reference model when communicating across agencies
- the current level of detail provided within the GSBPM

The major unresolved question identified in that feedback, however, was how the ABS (and others) can continue to engage with the ongoing evolution, extension and international application of the GSBPM. This is discussed further in Section VIII.

Annex 4 : Placing the GSBPM within ABS Enterprise Architecture

Various EA frameworks exist. The ABS now applies TOGAF (The Open Group Architecture Framework).

TOGAF (similarly to most other EA frameworks) recognises four broad architectural domains

- business
- applications
- data
- technology

While the ICT division leads the application of TOGAF within the ABS, the resulting ABS Enterprise Architecture is seen as belonging to the organisation as a whole. In order for this ownership to be meaningful, all divisions need to participate in defining, agreeing and supporting the realisation of relevant aspects of the architecture. For example, there is no point defining a business architecture that isn't expressed in a way that is meaningful to, agreed by and actionable for both statistical business areas and for those charged with supporting their statistical business processes in various ways (e.g. the statistical methods division and ICT division).

The ABS has an architectural preference for utilising and supporting open, community driven, standards such as SDMX, DDI and Open Document Format. The selection of TOGAF as an EA framework can be seen as a decision aligned with this preference. While not of the same magnitude, and not being a formally managed standard in the same sense, ABS interest in GSBPM can also be viewed as fitting this pattern.

As described in Annex 1, the “high level framework” (Caterpillar) developed by BSIP was often used more broadly as a reference model in the business domain by ABS EA prior to the emergence of the GSBPM. This usage occurred by default rather than as a result of a specific corporately endorsed decision. (Such a decision might have made it more likely that business areas outside the scope of BSIP would have at least applied the same reference model for the statistical business process.)

From V1.0 onwards the GSBPM was recognised as likely to add value as a reference model in the business domain.

There was immediately a question, however, about which area within the ABS should take the lead in regard to the GSBPM, such as co-ordinating feedback to the UNECE and promoting broader understanding and application of the GSBPM within the ABS.

- The statistical methods division had an interest in the GSBPM in the context for which it was originally developed within METIS, namely as the reference model for the statistical business process when describing relationships between the “data” domain (with a particular, but not exclusive, focus on statistical data and metadata in an NSI context) and the “business” domain

- The ICT division had an interest in the GSBPM from a broader architectural perspective

- “Business” divisions generally favoured their locally developed process models and, at least initially, did not place a priority on the GSBPM in regard to the “business” domain

As the statistical methods area had other connections with METIS it ended up taking the lead, working closely with the ICT division. If, however, the GSBPM were considered as an independent architectural “artefact”, without its connection with METIS, the lead role within the ABS may not be with the same area.

Annex 5 Relevance of the DDI Life Cycle Model

As highlighted throughout this paper, the ABS recognises great value in the GSBPM as a reference model when addressing statistical business processes within the ABS and with other NSIs (e.g. when sharing information, experiences and plans related to these processes). In terms of the level and nature of detail, the GSBPM tends to be much more suitable for such purposes than the DDI LCM.

The DDI LCM, however, supports perspectives that the GSBPM does not attempt to provide. It is particularly valuable when considering, at a high level, the relationship between providers and users of statistics. When discussed within the National Statistical Service (NSS) it provides a perspective that is less "introverted" than the GSBPM. It helps emphasise that a key driver from ABS interest in DDI (and SDMX) underpinning statistical processes in future are benefits to the operation of the NSS, not just the ABS.

The DDI LCM is also simple and powerful in expressing the concept that those users of statistics who draw together data (typically from multiple sources) and add analytical (and other) value to it can (and, ideally, should) contribute back to the pool of existing statistical data. This is a particularly valuable message in the NSS context.

This concept from the DDI LCM can also be relevant when thinking about statistical compilation and similar analytical activities within the ABS. Discovering, understanding and harnessing outputs from other statistical business processes provides a starting point for these activities, which then go on to produce further outputs in their own right.

The same scenario can be thought of in GSBPM terms. For example, many aspects of the "Collect" phase for National Accounts link to the cusp of Analyse and Disseminate for other statistical activities across the ABS that contribute data to the National Accounts. Once sub-processes are added to the picture, however, a GSBPM based illustration of the statistical activities that contribute to National Accounts, as well as National Accounts as a statistical activity in its own right, becomes messy.

The positioning of "Data Archiving" within the DDI LCM (given the DDI use of the term "archiving" can simply designate appropriate stewardship of content in situ) can also help clarify roles and responsibilities in such scenarios where, within a single NSI, outputs from one statistical activity are repurposed as an input to another.

References

[1] At the time of writing, in addition to the current version of the GSBPM (V4.0), earlier versions (namely V1.0, V2.0 & V3.1) are accessible via
http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model

[2] The quote is from "ESSNET – COMMON REFERENCE ARCHITECTURE (CORA) QUESTIONNAIRE" which has been sent to all NSIs within the ESS (European Statistical System) and a number of others. Outputs from CORA will be in the public domain but the working documents are not. Public domain resources in the meantime include

1. "CORA for ITDG", briefing to the Eurostat IT Directors' Group in October 2009 (more recent)
   http://www.slideshare.net/vaccaricarlo/cora-for-itdg

2. "COmmon Reference Architecture (CORA) ESSnet: history and next activities", paper and presentation delivered to Meeting on Management of Statistical Information Systems (MSIS) in May 2009 (more detailed)
   http://www.unece.org/stats/documents/2009.05.msis.htm

[3] Primarily the Sharing Advisory Board (SAB) whose mission is to "provide strategic direction to international work on the convergence of statistical business architectures..."
http://www1.unece.org/stat/platform/display/msis/Software+Sharing

[4] "Statistical Methods" and "ICT" are not the formal names assigned to the two divisions within the ABS. Their formal names are "Methodology and Data Management Division" and "Technology Services Division". Rather than using their longer formal names, or using ABS specific abbreviations, some concise terms that are (hopefully) generally recognised outside the ABS were used. More information about the structure of the ABS is available from the ABS website.
http://www.abs.gov.au/websitedbs/d3310114.nsf/89a5f3d8684682b6ca256de4002c809b/2565ffa0329c50f5ca2572c2001e542e!OpenDocument

[5] TOGAF 9 refers to Version 9 of The Open Group Architecture Framework.

- A simple overview is provided in Wikipedia
  http://en.wikipedia.org/wiki/The_Open_Group_Architecture_Framework

- Official documentation related to TOGAF 9 can be browsed, and downloaded from
  http://www.opengroup.org/architecture/togaf9-doc/arch/

[6] Documentation of V3.1 states that version *is intended as an input to the METIS Workshop to be held in Lisbon on 11-13 March 2009, **where it is intended that the GSBPM will be finalised.*** (emphasis added)

[7] http://www1.unece.org/stat/platform/display/metis/METIS-wiki

All references and cited URLs were valid as of 2010.02.17. Content of web pages may have been updated subsequent to citation.