

CONFERENCE OF EUROPEAN STATISTICIANS

For discussion and
recommendations

Meeting of the 2017/2018 Bureau
Helsinki (Finland), 14-15 February 2018

Item III (f) of the Provisional
Agenda

**COMMENTS ON THE PROGRESS REPORT OF THE HIGH LEVEL GROUP
FOR THE MODERNISATION OF OFFICIAL STATISTICS (HLG-MOS)**

Note by the Instituto Nacional de Estadística y Geografía, Mexico (INEGI)

This paper presents three strategies to further support the development and adoption of the UNECE's HLG-MOS models.

The Bureau is invited to discuss the outcomes and provide recommendations on the work under HLG-MOS to be undertaken in 2018.

I. INTRODUCTION

1. The recent development of data processing technologies has driven National Statistics Institutes (NSIs) to get up-to-date in the production of statistical information. To face these challenges, the High-Level Group for the Modernization of Official Statistics (HLG-MOS) of the United Nations Economic Commission for Europe (UNECE) has developed two strategies. The first is the exploration of potential new sources of information. The second is the development of models that support the standard-based modernization of official statistics. This note focuses on the latter.

2. The modernization of official statistics with an emphasis on standardized models facilitates the development and implementation of many aspects of statistical production: planning, controls, monitoring and quality evaluation, risk mitigation, automation, software development, incremental improvement and training. In a nutshell it promotes better quality and economies of scale.

3. The CES endorsement of the HLG-MOS models in 2017 was a great step forward. Likewise, the approval of the Executive Committee to continue HLG-MOS activities from 2018 to 2022 provides new opportunities but also new challenges that need to be considered. The most relevant are the provision of a general roadmap for the development of models and the adoption of a framework that supports a consistent development across different models.

4. The next section of this note describes the opportunities that INEGI has observed in the adoption of HLG-MOS models. The third section suggests three approaches that could help to address the observed opportunities. The last section summarizes the note.

II. BACKGROUND

5. INEGI is investing heavily in the adoption of HLG-MOS models. By implementing these models, our adoption task teams have identified some gaps in the current state of them:

- A roadmap of how HGL-MOS models will develop in the future. This is important as the NSIs need to figure out where to devote resources in case a model is not yet fully developed. For instance, what will be the timeline for the development of the overarching processes of the Generic Statistical Business Process Model (GSBPM). Thus far, there is development in few of the 15 overarching processes and the efforts do not make an explicit link. For instance, Data Management related activities are being developed under the Common Statistical Data Architecture (CSDA) umbrella.
- Organizational capabilities. NSIs should know which the minimum capabilities are in order to adopt specific models. For instance, the technical background needed to understand and implement a complex model such as Generic Statistical Information Model (GSIM). The Modernization Maturity Model (MMM) measures *ex-post* the level of adoption. However, it provides neither an *ex-ante* guide nor what is needed to move from one level to the next.
- Explicit subject matter on each model. There is a need to establish a baseline of knowledge for the task teams in charge of adopting the models.
- A guide to develop a roadmap of how an NSI could adopt more than one model.
- Consistency in the way the models are conceptualized. For instance, GSBPM is descriptive, but Common Statistical Production Architecture (CSPA) is prescriptive. This is, CSPA describes what is expected but does not provide an explicit model.
- A systematic set of guidelines to coordinate the adoption of models.
- Explicit description of interrelationships and dependencies among models.

6. The extent of the needs expressed above varies widely, particularly across different models and standards. GSBPM is more mature, has some associated models and has more examples of implementations. However, the rest of the models are still short on supporting documentation.

III. PROPOSALS

7. We encourage the HLG-MOS task teams to adopt three general strategies:

- To adopt a general framework to support the development of a roadmap for an integrated approach to all models.
- To explicitly incorporate frameworks and methodologies already developed by industry.
- To create training materials for all models to help NSIs on the adoption of HLG-MOS models.

8. Some of these approaches have been either implicitly adopted or explicitly adopted by HLG-MOS models but in a moderate form. We consider that a focused adoption on the three strategies could fill the gaps identified in the previous section.

A. Adoption of a general framework

9. The experience of INEGI adopting the HLG-MOS models has raised the need to have a general framework to include a holistic approach and to develop a long-term strategy. Most NSIs, INEGI included, have started adopting GSBPM. However, as soon as INEGI started formalizing the GSBPM's adoption the need to connect with other processes and domains became apparent.

10. Some HLG-MOS models make an implicit reference to [The Open Group Enterprise Architecture Framework](#) (TOGAF) to define the domain it pertains, that is Business, Data or Application Architecture. However, we suggest to adopt TOGAF not only to specify the architecture domain, but also to provide a general framework that can place each model in its appropriate context.

11. Enterprise Architecture is about managing complexity of the Business-IT landscape in a specific context. The three main reasons to engage with it are (Wierda, 2015, pp. 35-36):

- To prevent IT-chaos in a continuously changing landscape.
- To make sure the Business-IT landscape fits the strategic business goals.
- To make sure the business profits from opportunities enabled by IT.

12. These three reasons can be viewed as a progression in maturity. First you want to prevent chaos from emerging, then you want to change your IT landscape according to strategic business requirements and finally you want to be able to incorporate IT innovations to your business.

13. TOGAF is not only one of the most consolidated frameworks for enterprise architecture but has methods, guidelines and techniques that can support the whole HLG-MOS project. The most relevant are:

- **Governance.** To guide the overall model adoption from a senior executive standpoint.
- **Principles.** To provide guidelines to procure consistency among models.
- **Enterprise Continuum.** To explicitly identify the frameworks and methodologies available in industry that could be useful for NSIs.
- **Architectural Artifacts.** To formally document elements created to describe a system, solution, or state of an enterprise. For instance, GSBPM subprocess and GSIM objects.
- **Content Management.** To develop a container to classify reusable artifacts and the relationship among them.
- **Capability Maturity Models (CMMs).** To provide evaluation methods to gain control and improve the change processes.

- **Compliance.** To establish the degree of compliance of NSIs developments vs. HLG-MOS model specifications.
- **Architecture Development Model (ADM).** To define roadmaps for the development of models and NSI's adoption.

14. The goal of the HLG-MOS teams has been to develop a generic Enterprise Architecture to support standardization in the production of official statistics. However, this goal has not been made explicit. It is a good time to make it explicit so that a comprehensive framework can be introduced to define a roadmap for the short and long term work of the task teams. Table 1 is organized around TOGAF's architecture domains (TOGAF 9.1, [core concepts](#)) :

- **Business architecture.** Defines strategy, governance, organization and key business processes.
- **Data Architecture.** Describes the structure of the organization's data assets and data management resources.
- **Application Architecture.** Provides the blueprint for applications' deployment, interactions and relationships to the business processes.
- **Technology Architecture.** Describes the logical software and hardware capabilities that support the other three domains.

15. In Table 1, a description of the subject matter is provided, as well as some of the most notable industry practices¹. To facilitate implementation of an enterprise architecture, it is necessary to have a conceptual description and a model for all domains. In Table 1 a match is performed between the HLG-MOS models developed so far and TOGAF domains divided by the conceptual description and the model. This makes it easy to identify gaps in development for HLG-MOS models.

16. Table 1 shows the usefulness of adopting a general framework under which modernization of statistics through standards can be viewed as one entity. Assuming that general framework to be enterprise architecture as it is understood in TOGAF, there is a need for the development, adoption of industry practices or adaptation of industry practices on: the model for the application architecture; the conceptual description and model for the technological architecture. Furthermore, a revision of current standards should be made in order to guarantee their alignment to the adopted general framework.

17. Although strategy is country specific, it is the umbrella under which the modernization activities must be logically aligned. It is further suggested that Data management only needs a conceptual description and not a model. The specific strategy suggested to fill these gaps is further detailed in the next section.

18. It is worth mentioning that the term Enterprise Architecture is prone to create confusion with non-technical users. The approach used at INEGI is to place it under the umbrella of the Committee for Quality Assurance. The EA term is never used. A similar approach can be used at UNECE with the term ModernStats. As it is shown in Table 1,

¹ The reference to industry practices is not comprehensive.

strategy, governance and organization is country specific but are key elements to undertake a successful implementation of enterprise architecture.

Table 1.
TOGAF Domain relationship with HLG-MOS models

TOGAF Domain	Subject matter	Industry practices	Conceptual description	HLG-MOS Model
Business Architecture	Business strategy	OECD best practices, UN principles, Country's strategic plan.	Country specific	Country specific
	Governance			
	Organization			
	Key business processes	Business process specific	GSBPM	GSBPM
	Over-arching processes	DCAT, DMBOK, DCMM, DDI, ISO 9000, O-ISM3, O-RA, PMI, SDMX	Data Management: CSDA, Not yet developed for other over-arching processes	Not yet developed
Data Architecture	Organization of logical and physical data assets	Data Modeling	GSIM, LIM	GSIM
Application Architecture	Software Services	Service Oriented Architecture (SOA), ESS SPRA	Limited with CSPA	Not yet developed
Technology Architecture	IT architecture and services	IT4IT	Not yet developed	Not yet developed

19. Needless to say, the adoption of TOGAF will not solve all inconsistencies and coordination problems among domains and models but, by providing a general framework, the development of models could have a clearer path forward.

B. Adoption of industry practices and methodologies

20. Historically, industries have faced problems that have been solved. Solutions in one industry were in many cases reused by other industries. For instance, Internet was originally developed to interconnect academic and research institutions. The solution of this “industry” was eventually used by other industries around the world. An industry solution became a generic solution. Likewise, the Official Statistics “industry” can adopt solutions such as frameworks, methodologies and standards from other industries. Only when such solutions do not exist, should NSIs undertake developments from scratch.

21. The creation of GSBPM was a necessary development for Official Statistics, no such model existed. This model implicitly adopted a process-oriented approach developed in industrial engineering to document many years of world-wide experience on statistical production. Currently, GSBPM is not only a reference model, it is the documentation of best practices to produce official statistics and a *de-facto* standard.
22. Nevertheless, a revision of GSBPM is scheduled for 2018. NSIs are asking for further developments of overarching processes, the smoothing of the “survey flavor” of the paper and an explicit guide to interrelate with other models. Moreover, it seems the model can be used to support production of non-statistical information such as geographical information.
23. The development of GSIM provided a comprehensive conceptual data model. The adoption of GSIM has not been as widespread as GSBPM. A conceptual description of the model, implementation guidelines and examples of the interactions between GSBPM and GSIM are still needed. Likewise, the Logical Information model (LIM) is starting to define a possible jump from a conceptual to a logical model, but more work needs to be done.
24. GSBPM and GSIM had no equivalent in other industries but current work on other HLG-MOS models overlaps with models and definitions that have already been developed. For instance, the CSPA paper provides a set of recommendations for software component development that are a subset of the industry’s Service Oriented Architecture (SOA). The definitions for SOA are publicly documented on the site of the [Organization for the Advancement of Structured Information Standards \(OASIS\)](#) and [The Open Group](#) site. SOA recommendations are not only programming recommendations, they can help an institution to develop a comprehensive software development capability. Likewise, the current version of CSPA does not provide an architecture² because it does not have a set of application components. An example of application architecture that could be adopted is documented in the [ESS Statistical Production Reference Architecture](#).
25. The domain of Data Architecture in HLG-MOS models has already been defined in GSIM/LIM. The work of the group on Common Statistical Data Architecture (CSDA) has been around the development of Data Management definitions that are already documented as best practices in the [Data Management Body of Knowledge](#) developed by The [Data Management Association International](#) (DAMA). A similar context of Data best practices can be found in the Data Management Maturity Model (DMM) available at the [Capability Management Model Institute \(CMMI\)](#) and the Data Management Capability Assessment Model (DCAM) developed by [Enterprise Data Management Council](#).
26. The development of GAMS0 was a first step towards having a general framework. It was also a starting point for the introduction of Enterprise Architecture, such as TOGAF. However, as it was mentioned in the previous section, there are opportunities to adopt more tools and techniques available in TOGAF. The final objective being the creation of a Reference Architecture for Official Statistics.

² TOGAF defines and Architecture as: A formal description of a system, or a detailed plan of the system at component level to guide its implementation.

27. The Modernization Maturity Models (MMM) are a first step to introduce the concept of maturity and levels of adoption. However, the MMM does not allow for the identification of the dimensions that need to be developed to move capabilities to higher level of maturity. The HLG-MOS needs to move forward on this subject and the incorporation of industry models should have a thorough evaluation so that its adaptation can be useful in the official statistics' context.

28. The Capability Maturity Model Institute (CMMI) provides maturity models for measuring capabilities of:

- People
- Data Management
- Services
- Quality

29. Likewise, The Open Group offers references to evaluate maturity on:

- Software services
- Enterprise Architecture

30. The mentioned frameworks could support work on Data Integration, Data Architecture, Capability Definitions, Software Development, Documentation and the revision of GAMS0. We encourage not to release more models, or new versions of current models, until a reference to industry models and guidelines has been performed.

31. The list of models and frameworks delineated in this section are not meant to be exhaustive. Other models could and should be incorporated. It is fair mentioning that the proposals are generic industry models that can be tailored to specific NSI needs. In this context, the work of some of the HLG-MOS teams could move from developing models from scratch to provide guidelines for the adoption of industry best practices and models.

C. Development of training material

32. The work of the Task Group on Training for the GSBPM will be welcome. A similar approach will be needed to create material that supports the understanding of each model. The adoption of GSBPM and related models at INEGI has made evident the different versions of training materials that need to be developed for each of the current models. In a nutshell, there is a need to create documentation for:

- Subject matter - the reference to the basic theory that supports the model.
- Conceptual Model - A high level abstraction of the model.
- Logical Model - A formal description of the model.
- Physical model - An example of an implementation.

33. Examples of training materials which need to be provided in GSBPM:

- **Subject domain.** The general theory and practice that supports the activity of the model. For instance, in GSBPM there is a need for material explaining what is a process-oriented approach, process decomposition, formal modeling such as IDEF, BPM and the like.

- **Glossary.** Specific terms for the model that are also consistent with other models.
- **Generic Conceptual Model.** For instance, adaptation to local language, local regulation, laws and institutional restrictions.
- **Generic Logical Model.** Graphical models using techniques available in industry (Business Process Models (BPM), Archimate, Use Cases).
- **Generic Physical Model.** Hypothetical implementations (or actual implementations) that include enforcement and legal support of the actions stated in the model.

34. The material can include:

- Documents
- Presentations
- Workshops
- Videos
- Online courses
- Blogs and wikis

4. SUMMARY

35. The work developed by the HLG-MOS groups has been path-breaking and extremely useful for NSIs. However, as more models are available it is more challenging to keep all models consistent, and providing a general roadmap becomes not only convenient but necessary.

36. Current and future work should try to adopt existing frameworks and models so that the work of the task teams can concentrate on the adoption rather than on the development of the different domains. The adoption of industry models would facilitate both the hiring of staff with the required skills and the training of existing NSIs staff.

37. As models are incorporated to the HLG-MOS list, the use of a framework becomes more relevant. TOGAF can support the development of HLG-MOS models and the adoption of all models in NSIs.

38. The development of a document with the conceptual description of a model is a first step. To support the needs of NSIs, the development of a broad set of materials is necessary, such as references to the subject matter, specific training materials for the model, guides for use and implementation and, eventually, a methodology for the evaluation of success of an implementation.

LIST OF ABBREVIATIONS

ADM - Architecture Development Model
BPM - Business Process Models
CMM - Capability Maturity Models
CMMI - [Capability Management Model Institute](#)
CSPA - [Common Statistical Production Architecture](#)
CSDA - [Common Statistical Data Architecture](#)
DAMA - [Data Management Association International](#)
DCAM - Data Management Capability Assessment Model
DCMM - Data Capability Maturity Model
DMBOK - [Data Management Body of Knowledge](#)
EA - Enterprise Architecture
ESS SPRA - [European Statistical System's Statistical Production Reference Architecture](#)
GAMSO - [Generic Activity Model for Statistical Organizations](#)
GSBPM - [Generic Statistical Business Process Model](#)
GSIM - [Generic Statistical Information Model](#)
HLG-MOS - [High-level Group for the Modernisation of Official Statistics](#)
IDEF - [Integrated DEFinition Methods](#)
INEGI - Instituto Nacional de Estadística y Geografía (Mexico's NSI)
IT - Information Technology
IT4IT - [The Open Group reference architecture for managing the business of IT](#)
LIM - [Logical Information Model](#)
MMM - Modernization Maturity Model
NSI - National Statistical Institute
OASIS - [Organization for the Advancement of Structured Information Standards](#)
OECD - Organization for Economic Co-operation and Development
O-ISM3 - [The Open Group Information Security Management Maturity Model](#)
O-RA - [Open Risk Analysis](#)
SOA - Service Oriented Architecture
TOGAF - [The Open Group Enterprise Architecture Framework](#)
UN - United Nations
UNECE - United Nations Economic Commission for Europe

* * * * *