**Integrating process documentation and quality management for statistical production**

*Silvan Zammit, Director Data Capability, National Statistics Office,* *silvan.zammit@gov.mt*

*Anne Marie Pace, Head of Methodology and Research, National Statistics Office,* *anne-marie.pace@gov.mt*

**Abstract**

*With the aim of increasing efficiency and productivity, and, above all, improving the quality of its statistical output, the National Statistics Office (NSO) of Malta embarked on a three-year project for the development and expansion of a quality management function that covers technical issues and ensures business continuity. The main goal of this project is to integrate data and metadata standards based on the ESS Reference Metadata Reporting Standards and use them as a template for process documentation in line with the Generic Statistical Business Process Model. This would in turn provide a framework for process quality assessment and improvement.*

**Keywords:** quality management, harmonisation, GSBPM, metadata.

1. **Introduction**

In today’s ‘Information Age’, the demand for information has increased considerably and more than ever before. With the massive amounts of data that are generated as we speak, statistical offices around the globe face a dilemma as they are constantly required to produce better quality statistics in less time.

While the heterogeneity of data available for statistical production offers a unique opportunity for statistical agencies to bring together many different data sources to their advantage - from the most traditional acquired through personal interviews to the most sophisticated generated by physical devices - this in itself poses an additional challenge.

In turn, accurate statistical information necessitates efficient and harmonised production in line with established standards and methodologies to promote the re-usability of statistical products and services.

Taken together with the need for timely statistical information that focuses on its users, this requires statistical agencies to position themselves as the main agents to compile statistical data and make it available to a wider audience in line with international standards. With this in mind, the National Statistics Office (NSO) of Malta launched a three-year project which sets quality enhancement actions targeted at different statistical processes and user groups. The integration of harmonised metadata standards and quality management are high on the agenda.

1. **The national context**

Before the NSO embarked on this project, growing national and international requirements and their fulfilment were always considered as priority areas, although the situation was suboptimal due to lack of internal coordination and efficiency. With the variety of tailored programs used across domains for data collection and data mining, a number of non-harmonised processes adopting customised metadata were created over time, which led to an inability to optimise data consumption by other domains and external users.

In view of these issues, the NSO set a target to improve the quality of its output and become more efficient and relevant to its users. In times of diminishing resources this could only be achieved by improving internal data management practices through established and harmonised internal processes aimed at maximising data currency.

1. **Integrating data and metadata standards**

Standards facilitate everyday life. They ensure that products, services and methods are appropriate for their intended use. Intentionally or not, we use standards in all aspects of our daily lives – communications, media, healthcare, food, transport, construction, furniture, energy – and statistics is no exception.

Standardisation in statistical research safeguards the interoperability between processes for various statistical domains, and ensures that statistical products are compatible across such domains. Taking these aspects into account, the NSO identified the need for standardised documentation of all activities that make up the entire statistical business chain, which was previously almost inexistent.

The most logical starting point was to get to know *ourselves* better. This would allow the identification of any areas of improvement and possible ‘black-box’ systems, which tend to be created in extensive processes involving large groups of people.

To portray the NSO’s statistical processes coherently and allow for comparability across distinct statistical domains, the Generic Statistical Business Process Model (GSBPM)[[1]](#footnote-1) was adopted in its original form.

The GSBPM was foreseen as the ideal supporting framework for process mapping and output metrics evaluation. It had also been tried, tested and improved by other statistical agencies who adopted it for similar reasons over the years. In conjunction with the development of a quality management framework, it would allow the assessment of quality at each phase of the statistical production chain, and the identification of specific services that could possibly be re-proposed in other domains.

Statistics that are reconcilable require a set of common concepts and classifications, in line with requisites of methodological manuals and regulating frameworks which are applied internationally. This has highlighted to the NSO the need for embedding a metadata repository in a dedicated portal[[2]](#footnote-2) that satisfies the needs of various levels of users.

The harmonisation of metadata in line with ESS Reference Metadata reporting standards entailed a mammoth task: the scrutiny of all concepts, definitions and classifications covered by all statistical domains. The Single Integrated Metadata Structure (SIMS), a dynamic inventory of quality and metadata statistical concepts aiming at documenting methodologies, quality and statistical production, was considered. SIMS may be assumed to be a fusion of the Euro-SDMX-Metadata Structure (ESMS) and ESS Standard for Quality Reports Structure (ESQRS) reference metadata standards, available at ESS level.

Following a thorough evaluation of the terminology relative to a set of key methodological manuals available at national and European level, more than 400 distinct concepts and 40 standard classifications were included in the metadata portal.

In addition to providing a standard definition of each term, all statistical concepts and definitions were presented alongside a list of related variables and existing data sources to better facilitate and enhance user experience. The list is regularly updated over time to reflect newly emerging statistical areas and products to assist all levels of users in their interpretation of statistical outputs.

An important facet of the NSO metadata portal, aimed at users who wish to get a better understanding of main methodological and quality-related issues, is the availability of reference metadata. Indeed, specific reports for 45 main statistical domains are documented in line with the ESMS or the ESQRS metadata standards.

These reports are updated annually and are intended to provide a more complete coverage on methodological and quality-related issues, as well as to complement basic quality indicators available on the NSO website.

1. **Introducing standard documentation for work processes**

Process documentation is vital for efficient business operations. It facilitates business continuity and operational redundancy, business process revaluation and improvement, in addition to the implementation of risk and quality management practices. It may be considered as an operational blueprint for an organisation, irrespective of the activity area within which it operates.

The documentation process started in 2015 following an internal study which investigated the level of interdependencies between domains. This allowed the identification of certain key areas, 33 in total, for which process-metadata documentation was required.

To facilitate the transition and maximise user acceptance and engagement, a series of consultation sessions and internal workshops with all relevant domain units were organised. Keeping in mind that more often than not users tend to be output-oriented, these sessions aimed at explaining in detail the need for having standard process documentation in place, the various uses of the GSBPM and expected benefits.

In addition, a comprehensive step-by-step guide on the GSBPM was provided to each internal user, highlighting the particular characteristics of the model at every level, guidelines on how best to document the various phases to ensure compliance with the model itself and the level of detail required under each sub-process.

Reports were drafted by the respective domain units and channelled through the Methodology department to ensure cross-domain consistency and a sufficiently detailed outline which is user-oriented. Following a rigorous vetting process, the reports were revised by the responsible domain units. Throughout the process, reports were cross-verified to ensure perfect synergy between domains in terms of data provision, data sources, methodology and classifications used.

Essentially, the idea was to collect detailed information about national and international requirements and obligations, details about all processes, deliverables, key players involved in various stages of each process, how each domain interacts with others, and the dynamics of such interaction, together with users’ needs.

In addition to the operational value of having adequate process documentation in place, this phase of the project also allowed the mapping of certain processes including any existing interdependencies, and the evaluation of each. Operational and technical issues were addressed collectively by domain experts, methodologists and IT people.

Once a comprehensive set of reports was compiled, an internal repository for the storage of process-metadata was created on the NSO intranet. A holistic approach was taken to the repository, including reports for all domains, technical annexes and supporting documentation, legislation, IT requirements and methodological documents. NSO employees are constantly encouraged to refer to these reports not only to facilitate day-to-day operations but also as an effective knowledge management tool.

1. **A framework for quality management**

Integrated in the GSBPM in form of an overarching process[[3]](#footnote-3), quality management, in its simplest form, describes and defines quality assessment and control mechanisms that take place within a typical statistical organisation. Such a framework extends and complements the GSBPM by adding additional activities needed to support statistical production to assess the fitness-for-purpose of the links that exist between various sub-processes in the GSBPM.

At its core, the application of a Quality Management Function (QMF) at the NSO has the objective of ensuring quality in specific statistical processes through the application of standard methodologies, metadata and practices. It has been developed keeping users of statistical data at its centre: internal staff to facilitate the integration of various data sources in a unified approach, and external users so that these can understand better the complete life cycle of the data they are using.

Following a detailed account of the literature available on quality management at international level, the NSO opted to focus on compliance with the European Statistics Code of Practice (CoP[[4]](#footnote-4)). Consequently, the CoP was the point of departure. Nonetheless, relevant aspects of national law, such as provisions of the Malta Statistics Authority Act[[5]](#footnote-5) which governs operations of the NSO were considered.

This phase comprised a course of action to ensure consistency with the CoP, covering 23 different topics, ranging from purely methodological to knowledge management and operational issues.

The topics are presented in separate chapters, each with an opening section on the regulatory framework within which the set of relevant statistics is produced. This section is very often complemented with the terms of the relevant principle or indicator of the CoP. Besides, the legal frameworks in which the obligations are prescribed are duly referred to in every section.

Apart from furnishing statisticians with guidelines to facilitate the correct execution of statistical procedures, this document is also intended to serve as a handbook for personnel training. With this in mind, each chapter comprises a section containing clear definitions of applicable principles. The third and last section of each chapter is, in essence, the central component that provides quality guidelines and suggested best practices.

As any other statistical agency, the success of the NSO largely depends on the extent to which it can meet user requirements and stimulate their appetite for information. Fulfilling this role calls for a healthy collaborative environment not only within the organisation but also with the data suppliers, based on trustworthiness and professional attitude. This concept evidences the need for making all statisticians and other workers aware of their responsibility to be fully committed to ensuring quality in their work.

In view of the importance that administrative and other forms of auxiliary data have acquired in today’s statistics, a significant part of the QMF is dedicated to the access to and proper use of administrative data. The role of administrative data in modern statistics cannot be emphasised enough, as it allows statistical agencies to improve the quality of their work while minimising administrative costs and burden on data providers.

To maximise the uptake of administrative data, the harmonised metadata practices outlined above were extended to secondary data sources, covering not only the concepts and definitions employed, but also cooperation agreements and other issues of an administrative nature.

Having a more direct and open relationship with data subjects ensures that statistical agencies remain relevant in the eyes of their users and create a demand-driven environment. For this reason, user needs are addressed in detail in the document with specific reference to confidentiality aspects and the need for clarity and simplicity in data collection and dissemination tools. The QMF also stipulates the need for maintaining professional independence, impartiality and transparency at all times.

Consequently, procedures for careful recruitment, training, and assistance to field officers are also covered. In addition to the importance of having a courteous approach towards respondents, the reliability of interviewers to acquire precise information was manifested in the QMF as a crucial element for the successful collection of data.

Stemming from the NSO’s commitment to apply sound methodologies in its processes, a wide range of technical topics were addressed. Amongst others, these include guidelines on statistical editing and imputation, construction of indices, weighting methodologies and consistency checks. Good practices were recommended for distinct scenarios which statisticians are confronted with in the execution of their tasks. The varied needs of different statistical areas necessitated comprehensive overviews of each topic, providing clear guidance to all readers.

From a more qualitative and less technical perspective, the QMF includes extensive guidelines and best practices on adequate metadata documentation and quality reporting. The main reason for embedding metadata aspects in the QMF was to highlight its importance in various phases of statistical production.

Statistical agencies are made up of people with different statistical knowledge bases and degrees of belief. For this reason, metadata is at times overlooked rather than considered as an integral part of every process.

Nowadays metadata has indeed become the cornerstone of every statistical work, which must be thoroughly evaluated and understood before any data linkage or statistical interpretation can be made.

**Fig 1. NSO’s Architectural Framework for Data and Quality Management**

In this context, the set of recommendations proposed in the QMF lay emphasis on the need to regard metadata documentation as an integral part of the statistical production process across the organisation. Standard documentation systems were described in detail and the context for when metadata or quality documentation are required was explained.

1. **Using standards to assess quality: the link between QMF and GSBPM**

Departing from the primary scope of the GSBPM model to standardise the terminology of statistical processes, the model presented above contributes towards a systematic approach for the recognition of good quality practices by confronting various processes. Having statistical domains documented in a standardised fashion facilitates mechanisms to assess quality (UNECE).

The QMF developed by the NSO may be regarded as a tool for bringing quality to output and process levels. To improve quality, it is essential to make quality assurance an integral part of the processes, as argued above. Essentially the nature of the GSBPM model and the QMF requires the interoperability of the two, resulting in a hybrid system that allows constant re-evaluation of internal work processes, methodologies and knowledge management to enable the Office to keep pace with technological and methodological advances.

The main line of thought here is that methods for assessing quality in an all-encompassing fashion cannot be detached from a thorough evaluation of statistical business process metadata in GSBPM form. This calls for an organisation that keeps assessing itself over time in an attempt to break down methodological statistical processes into small building blocks that may be re-used not only internally but also offered to the external world. This would also facilitate quality assessment at various phases, which in their turn can be re-used for various domains.

1. **Conclusion**

Ultimately, this is the point of convergence of data, process and knowledge management. Quality assessment at the NSO is now the result of the ongoing evaluation of process metadata. This evaluation allows the NSO to outline action plans that focus on various aspects, with special attention given to methodological soundness and harmonised metadata.

The review of business process documentation was a key factor, making it possible to identify any improvement actions that would ensure full compliance with the CoP and international standards, in addition to being more effective and consequently more productive.

1. **References**

Organisation for Economic Co-operation and Development (2018) *GSBPM v5.0*. Available at: [https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0](https://statswiki.unece.org/display/GSBPM/GSBPM%2Bv5.0) (Accessed 5 May 2018).

Eurostat (2018) *ESS Reference Metadata Reporting Standards*. Available at: <http://ec.europa.eu/eurostat/data/metadata/metadata-structure> (Accessed 6 May 2018).

Eurostat (2015) *Single Integrated Metadata Structure (SIMS) v2.0 and its underlying reporting structures: the ESS Quality and Reference Metadata Reporting
Standards ESMS 2.0 and ESQRS 2.0*. Available at: <http://ec.europa.eu/eurostat/documents/64157/4373903/SIMS-2-0-Revised-standards-November-2015-ESSC-final.pdf/47c0b80d-0e19-4777-8f9e-28f89f82ce18> (Accessed 4 May 2018).

Eurostat (2014) *Technical Manual of the* *Single Integrated Metadata Structure (SIMS)*. Available at: <http://ec.europa.eu/eurostat/documents/64157/4373903/03-Single-Integrated-Metadata-Structure-and-its-Technical-Manual.pdf/6013a162-e8e2-4a8a-8219-83e3318cbb39>. (Accessed 4 May 2018).

1. The GSBPM ([https://statswiki.unece.org/display/GSBPM/GSBPM+v5.0](https://statswiki.unece.org/display/GSBPM/GSBPM%2Bv5.0)) provides a standard framework and harmonised terminology to help statistical organisations describe their statistical production processes, as well as to share methods and components. It can also be used for integrating data and metadata standards, as a template for process documentation and to provide a framework for process quality assessment and improvement (UNECE). [↑](#footnote-ref-1)
2. The NSO metadata portal may be accessed at: <https://nso.gov.mt/metadata/>. [↑](#footnote-ref-2)
3. This is relative to GSBPM version 5.0. In future versions of the model, the Generic Activity Model for Statistical Organizations (GAMSO) is expected to expand over the current version of the GSBPM by including additional activities needed to support statistical production. No direct reference is made to GAMSO here since its applicability in the Maltese context will be assessed in due course. [↑](#footnote-ref-3)
4. <http://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice>. [↑](#footnote-ref-4)
5. <https://nso.gov.mt/en/nso/About_NSO/Documents/Legislation/MaltaStatisticsAuthorityAct.pdf>. [↑](#footnote-ref-5)