Statistics South Africa (Stats SA)

UN Expert Group on national Quality Assurance Frameworks (NQAF) – answers to initial questions

Marlize Pistorius

1. What are your country's or agency's experiences – if any – with the development or use of an NQAF?

   Stats SA do not have a Quality Assurance framework but we have a Quality Assessment Framework called South African Statistical Quality Assessment Framework (SASQAF) and we also published SASQAF. In 2006 a policy was developed called "Data Quality Policy 001: Policy on Informing Users of Data Quality, where the purpose of the policy was to ensure that producers of statistics in Stats SA adhered to the agreed standards, procedures and guidelines on producing quality data. Furthermore, the policy serves to inform users about the quality of data and products disseminated by Stats SA.

   In 2008 a working group was established to enhance the South African Statistical Quality Assessment Framework (SASQAF) and publish SASQAF edition 2. Along with SASQAF a generic set of operational standards and guidelines were developed. Stats SA also adopted the Statistics Canada's Quality Guidelines and as a secondary source we used Stats Finland's Quality guidelines handbook. We adapted these manuals for Economic Statistics and Household–based statistics called "Generic Operational Manual for Economic Statistics" and "Generic Operational Manual for Social and Population Surveys". Stats SA adopted the Generic Statistical Business Model (GSBM) and made small amendments to allow localisation (see appendix A). Stats SA mapped the GSBM to the phase, sub-processes, quality dimension and quality indicators (see appendix B).

   Stats SA are using 8 dimensions of quality and for external National Statistics Services (NSS) partners use 9 dimensions of quality for assessment namely:

   - Relevance;
   - Accuracy;
   - Timeliness;
   - Accessibility;
   - Interpretability;
   - Comparability and Coherence;
   - Methodological soundness; and
   - Integrity
   - The 9th external dimension used is pre-requisites of quality.

   As can be seen in appendix C, the process of declaring quality as part of official statistics is firstly a quality declaration, then pre-selection of quality indicators, DQAT review, a report out of the DQAT review and a quality improvement plan.

   Stats SA also developed a web-based tool (a quality toolkit was also developed to assist NSS partners) to assist with the review of all statistics within the NSS which provide management with a dashboard of all dimensions and the quality of the indicators.

2. What problems and obstacles have you experienced or anticipate experiencing in developing and implementing an NQAF?

   The problems we experienced were that the development of SASQAF edition 2 and SASQAF operational standards and guidelines took much longer as was anticipates and it was resource intensive. One of the main problems we experienced is the lack of a quality culture and buy-in from senior management. No quality initiative can succeed without the buy-in from senior management. One of the problems as well is once we started development we were getting entrenched in the development and did not communicate often enough with stakeholders thus practical implementation took to long to happen and it is my believe similar to Norway that quality initiatives are sometimes theoretical and not practical enough.

   One of the challenges of Stats SA which is hampering the implementation of SASQAF is the fact that we do not have a dedicated team of quality experts appointed to be responsible for the training. A quality strategy

has been developed but due to resource constraints the rollout was stopped in order to review how we could rollout the use and implementation of SASQAF as well as review all statistical releases.

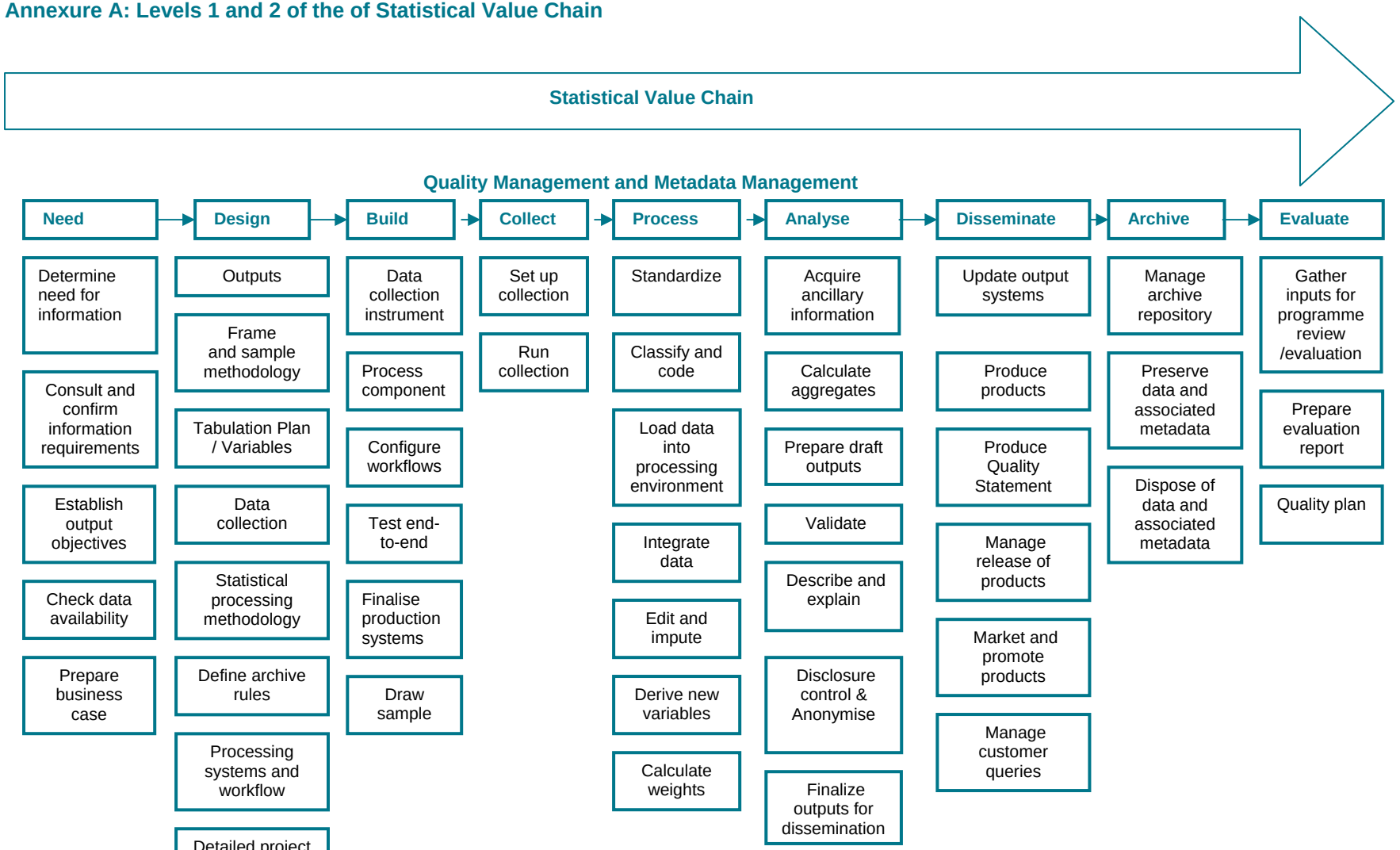
3. What are the main needs and priorities from a country perspective vis-à-vis the development and implementation of an NQAF?

   - The next step would be to roll out training initiatives in terms of quality training in SASQAF by having the necessary resources on board. Once teams of core Data Quality Assessment teams (DQAT) exists then Stats SA's will rollout the assessment and reviews of all statistical publications. The main need from Stats SA would not be how to develop standards, guidelines etc but the implementation of detailed standards, guidelines etc.
   - Agreed standards and definitions in terms of Quality Management.
   - International training initiatives on quality management.
   - Assistance with the drafting a quality management system for the organisation and the implementation there-off.

4. What are your initial comments regarding the three proposed templates for a generic NQAF presented in the Statistics Canada report http://unstats.un.org/unsd/statcom/doc10/2010-2-NQAF-E.pdf? (All three templates share a basic structure but differ in the way in which quality assurance procedures are presented).

   The document as compiled by Stats Canda is well written. I support that a template would probably be the easiest way to ensure comparability and consistency between NSO's but I believe it is important to map the quality Assurance frameworks also to the statistical value chain since guidelines and standards are written at that level. Secondly I'm also of the opinion that we might need to look at a fourth model which is a combination of different agencies and agree on a set of internationally best practices which can be sued as guidelines within agencies.

Annexure A: Levels 1 and 2 of the of Statistical Value Chain



Annexure B: Mapping quality indicators to activities in the Statistical Value Chain (SVC)

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Need	Determine need for information	Prerequisites of quality	1.1 The responsibility for producing statistics is clearly specified.
Need	Determine need for information	Prerequisites of quality	1.2 Standards and policies are in place to promote consistency of methods and results.
Need	Establish output objectives	Prerequisites of quality	1.3 Data sharing and coordination among data-producing agencies are clearly specified.
Need	Establish output objectives	Prerequisites of quality	1.4 Measures are in place to ensure that individual data are kept confidential, and used for statistical purposes only.
Need	Determine need for information	Relevance	2.1 Have both the internal and external users of the data been identified?
Need	Determine need for information	Relevance	2.2 Is there a process to identify user needs?
Need	Establish output objectives	Relevance	2.3 Are user needs and the usage of statistical information analysed?
Need	Establish output objectives	Relevance	2.4 Changes are made as a result of user needs assessments.
Need	Check data availability	Relevance	2.5 To what extent is the primary data appropriate for the statistical product produced?
Need	Prepare business case	Timeliness	4.4 Periodicity of release.
Need	Check data availability	Accessibility	5.1 Legal arrangements are in place to allow access to administrative records via manual, automated or electronic processes.
Need	Frame and sample methodology, Data collection	Methodological soundness	8.1 The scope of the study is consistent with accepted standards, guidelines or good practices.
Need	Check data availability	Integrity	9.5 Choice of source data, techniques and dissemination decisions are informed solely by statistical considerations.

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Design	Detailed project plan	Prerequisites of quality	Resources are commensurate with the needs of statistical programmes. • Staff • Facilities • Computing resources • Financing
Design	Detailed project plan	Prerequisites of quality	1.7 Measures to ensure efficient use of resources in 1.6 are implemented.
Design	Frame and sample methodology	Accuracy	3.5 Register/frame maintenance procedures are adequate. • updates. • quality assurance ▪ data audit.
Design	Data collection	Accuracy	3.6 Are data collection systems sufficiently open and flexible to cater
Design	Data collection	Accuracy	3.7 Description of record-matching methods and techniques used on the administrative data sources. • match rate as a percentage of total records • measure of false negative matches (missed matches) • measure of false positive matches (mismatches)
Design	Outputs, Frame and sample methodology, Tabulation plan/variables, Data collection Statistical processing methodology, Define archive rules,	Comparability and coherence	7.3 Data across comparable series, or source data are based on common frames, identifiers, concepts and definitions, and classifications, and departures from these are identified in the metadata.
Design	Tabulation plan/variables; Identify concepts	Methodological soundness	8.1 Concepts, definitions, and classifications used follow accepted standards, guidelines or good practices (national, international, peer-agreed).

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Design	Statistical processing methodology	Methodological soundness	8.2 Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: • questionnaire design • sampling • sampling frame design • frame maintenance • piloting • data collection • editing and imputation of data • analysis of data • revision data
Design	Tabulation plan/variables Identify concepts	Methodological soundness	8.1 Concepts, definitions, and classifications used follow accepted standards, guidelines or good practices (national, international, peer-agreed)
Design	Statistical processing methodology	Methodological soundness	8.3 Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: • questionnaire design • sampling • sampling frame design • frame maintenance • piloting
Process	Integrate data	Comparability and coherence	7.4 A common set of identifiers (for the purpose of record matching) exist and have been agreed upon by data producers.
Analyse	Validate	Accuracy	3.1 Measures of sampling errors for key variables are calculated. Amongst others these are:

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
			• standard error • coefficient of variation (CV) • confidence interval (CI) • mean square error (MSE) • design effect (DEFF)
Analyse	Validate	Accuracy	3.2 Measures of non-sampling errors are calculated, viz.: • Frame coverage errors • Systematic errors • Measurement errors • Processing errors • Model assumption errors • Non-response errors
Analyse	Define archive rules, Processing systems and workflow Validate	Comparability and coherence	7.2 Statistics are consistent or reconcilable over time
Analyse	Validate	Comparability and coherence	7.3 Statistics are checked for consistency with those obtained through other data sources.
Disseminate	Produce quality statement	Accuracy	3.3 Data from the primary source have been quality assessed..
Disseminate	Manage release of products	Timeliness	4.1 Average time between the end of reference period and the date of the preliminary results.
Disseminate	Manage release of products	Timeliness	4.2 Average time between the end of reference period and the date of the final results.

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Disseminate	Manage release of products	Accessibility	5.1 Are statistical products (e.g. data, metadata) available to the public?
Disseminate	Manage release of products	Accessibility	5.2 Rules governing the restricted availability of administrative records are well described and documented.
Disseminate	Produce dissemination products	Accessibility	5.3 Types of media and/or channels used for sharing data amongst stakeholders are adequate and preserve confidentiality.
Disseminate	Produce dissemination products	Accessibility	5.4 Data is accessible in a format beyond the producing agency.
Disseminate	Manage release of products	Accessibility	5.5 Statistics are released on a pre-announced schedule.
Disseminate	Manage release of products	Accessibility	5.6 Statistical products are made available to all users at the same time.
Disseminate	Manage customer queries	Accessibility	5.7 Statistics/administrative records not routinely disseminated are made available upon request.
Disseminate	Market and promote products	Accessibility	5.8 User support services exist and are widely publicised.
Disseminate	Manage release of products	Accessibility	5.9 Does a data dissemination policy exist, and is it accessible?
Disseminate	Manage release of products	Accessibility	5.10 Does the pricing policy governing dissemination exist, and is it accessible?
Disseminate	Market and promote products	Accessibility	5.11 Catalogues of publications and other services are available to users of statistics.
Disseminate	Manage release of products	Accessibility	5.12 Metadata are readily accessible to users.
Disseminate	Produce dissemination products	Interpretability	6.1 Documented metadata (definitional, operational, methodological, system and dataset) are sufficient to understand data.
Disseminate	Produce dissemination products	Interpretability	6.2 Statistics are presented in a clear and understandable manner. Statistical releases contain a summary of the key findings.
Disseminate	Produce dissemination products	Interpretability	6.3 Statistical releases contain a summary of the key findings.
Disseminate	Processing systems and workflow	Methodological soundness	8.4 Are revisions schedule followed? Are they regular and transparent?

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Disseminate	Manage release of products	Methodological soundness	8.5 Preliminary and revised data are identified in the metadata.
Disseminate	Manage release of products	Methodological soundness	8.6 Studies of revisions and their findings are made public.
Disseminate	Manage release of products	Integrity	9.1 The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated are available to the public and follow the UN principles of official statistics.
Disseminate	Produce quality statement	Integrity	9.2 Describe the conditions under which policy-makers, specifically government, may have access to data before release. Are the conditions published?
Disseminate	Produce quality statement Check data availability	Integrity	9.3 Advance notice is given of major changes in methodology and source data.
Disseminate	Outputs, Frame and sample methodology, Tabulation plan, Data collection Statistical processing methodology, Define archive rules	Integrity	9.4 Government commentary, when data are released, should be identified as such, and not be seen as part of the official statistics.
Evaluation	Gather inputs for programme review/evaluation	Prerequisites of quality	1.5 Measures to oblige response are ensured through law.
Evaluation	Gather inputs for programme review/evaluation	Relevance	2.6 Is there a process to determine the satisfaction of users with the statistical information?

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
All phases	All phases	Timeliness	4.3 Production activities within the statistical value chain are within the planned timelines, viz.: • data collection • data processing • data analysis • dissemination
All phase	All phases	Integrity	9.5 Ethical guidelines for staff behaviour are in place and are well known to the staff.

Annexure C: Procedure for designating statistics from organs of state as official statistics

1. The Statistician-General will publish and regularly update SASQAF as a framework within producing agencies who may apply to have data designated as official statistics. The Statistician-General, in consultation with the head of the producing organ of state or agency, determines the elements or outputs of the producing organ of state or agency to be designated as official statistics. These could include a survey, a register, a dataset, indicators, a data table, etc.

2. An organ of state or agency will apply, through the division responsible for the NSS at Statistics South Africa, to the Statistician-General to have their statistics designated as official statistics.

3. Applications will be referred to a Data Quality Assessment Team (DQAT) constituted by the Statistician-General, drawn from
   - Statistics South Africa
   - Applicant (organ of state)
   - Subject-matter expert(s) (recommended by the organ of state and/or the Statistician General)
   - Statistics Council member (observer status).

4. Appointed DQAT members will sign Terms of References for the review and an Oath of Confidentiality document.

5. For assessment to begin, the submitting organ of state and the statistics under review need to comply with three initial criteria:
   - The producing agency must be a member of the NSS.
   - The statistics are used to meet user needs beyond those specific and internal to the producing agency.
   - The statistics produced should be part of a sustainable series, not a once-off collection.
6. DQAT will assess the quality of the product(s) in terms of SASQAF requirements, assigning a SASQAF quality level to the product. The assessment process is as follows:

   a. the applicant will identify all the SASQAF indicators that are relevant to the product under evaluation, and motivate why the remaining indicators are not relevant;

   Note: The selection is based on the requirements of the product or the properties of the data. In principle, indicators that provide useful information to users should be selected. Not all indicators are relevant for all products.

   b. once DQAT and the applicant reach agreement on which indicators are relevant, and on the standard for each indicator, they will sign an agreement to this effect;

c. the applicant will then be asked to produce a quality declaration for their product, for all the agreed indicators; and

d. DQAT will assess these quality statements against the relevant standards, and based on the results, assign one of the four quality levels (quality, acceptable, questionable or poor), and will identify areas of improvement in the quality statements.

7. DQAT will recommend the overall SASQAF level of the product.

8. If the product submitted for evaluation is not classified as *quality statistics* in terms of the SASQAF levels of evaluation, DQAT will advise the applicant on areas of improvement.

9. If the product satisfies the requirements o*f quality statistic*s set out in SASQAF, the Statistician General will designate the product as *official statistics.*

10. Once the product has been designated as official statistics, it will be published with the Statistician-General's official seal of approval (the Official Statistics Mark), and stored in the NSS archive for public access.

11. The Statistician General will issue a notice in the government gazette to the effect that a product has been designated as official statistics

12. The product then becomes subject to periodic reviews, determined by the Statistician General in consultation with the head of the producing agency or department.

13. The Statistician General will publish the results of the assessment or review for access by the public.