12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Background document

Document INF/5-E
19 November 2014

English

SOURCE: ITU

TITLE: Data Quality Assessment Framework for ITU

International Telecommunication Union

Data Quality Assessment Framework for ITU

Data Quality Assessment Framework for ITU

1. Introduction

Several other international organisations, including OECD, ECB, and Eurostat have defined a *quality assurance framework* for their internal use in collecting, processing and disseminating statistical data. Although ITU has a very much smaller statistical program than these organisations there is every reason to suppose that such a framework, scaled appropriately to the scale of ITU's statistical activities, would be equally useful.

As the main mission of ITU is "to enable and foster the growth and sustained development of telecommunication networks and services…", the framework is referred to as a *data quality assessment framework (DQAF)* to emphasise that it refers to activities related to the collection processing and dissemination of statistical data, not to ITU as a whole.

The main benefits expected from the DQAF are that it will:

- provide a systematic mechanism for facilitating the ongoing identification of quality problems and possible actions for their resolution;
- provide a basis for creating and maintaining a data quality culture within ITU;
- stimulate and maximize the interaction among ITU staff involved in production or use of statistics;
- give greater transparency to the processes by which statistics are produced and their quality is assured and thereby reinforce ITU's image as a trustworthy provider of good quality statistics;
- provide reference material that can be helpful for training;
- provide a mechanism for the exchange of ideas on quality assurance with other producers and users of statistics, at international and national levels.

Whilst the DQAF was developed with the statistical operations of ICT Data and Statistics Division in mind, it is equally applicable to operations of other statistics producing units within the ITU, specifically the Regulatory and Market Environment Division (RME).

The intended readership/users of the DQAF are:

- IDS staff – the DQAF provides a framework for assessment of statistical activities;
- ITU and BDT senior management – the DQAF provides an indication how quality may be assessed;
- NSOs and NRAs – the DQAF provides quality guidelines and an indication of roles NSOs and NRAs as data providers can play in quality assurance;
- Data users – the DQAF provides users of ICT statistics with evidence of quality assurance by ITU.

Development of the DQAF

The DQAF was prepared by an expert in statistical data quality, Michael Colledge under the supervision and with the support of the ICT Data and Statistics Division. A preliminary outline

was discussed at a combined meeting of the Expert Group on Telecommunication/ICT Indicators (EGTI) and the Expert Group on ICT Household Indicators (EGH). Comments were obtained from members of both working groups and were taken into account in preparing this version.

In developing the DQAF there was no need to reinvent the wheel as several quality assurance frameworks for statistical organisations have been developed in recent years and advantage could be taken of their contents and experience in their use. Particularly influential international documents include the following.

- *Fundamental Principles of Official Statistics* – indicating how national statistical systems should be organized in order to produce appropriate and reliable data that adhere to appropriate professional and scientific standards.
- *Principles Governing International Statistical Activities* – comprising principles and practices that were developed and publicized by the CCSA and that should underpin the production of statistics by an international organisation;
- *UN Statistical Commission National Quality Assurance Framework (NQAF)* – a template that provides the general structure within which a country can formulate and operationalize a national quality assurance framework.
- *European Statistics Code of Practice (ESCoP)* – developed by Eurostat, comprising principles and indicators relating to statistical environment, processes and outputs of European NSOs and agencies.
- *European Statistical System Quality Assurance Framework (ESS QAF)* – developed by Eurostat, the focus of the framework is to assist in implementation of the ESCoP by European NSOs and Eurostat.
- *Data Quality Assessment Framework (IMF DQAF)* – developed by the IMF Statistics Division for use by NSOs and other national government agencies collecting and disseminating statistics.
- *Quality Framework and Guidelines for OECD Statistical Activities (OECD QFG)* – developed by the OECD for managing quality within its own organization.
- *European Central Bank Statistics Quality Framework (ECB SQF)* - developed by the ECB for managing quality within its own organisation.

As the organisations authoring these documents have much bigger statistical programmes than the ITU, although the quality principles remain much the same whatever the organisation, the framework itself has been appropriately simplified for ITU.

DQAF Components

The DQAF has four components.

- The first component is a set of *underlying principles*, providing the basis for formulating the DQAF.
- The second component is a set of *quality dimensions*, highlighting the various aspects of data and process quality.
- The third component is a set of *quality guidelines*, comprising good practices for assuring quality.

- The fourth component is *quality assessment and improvement program*, comprising a set of procedures for ensuring that quality is regularly assessed and appropriate quality improvement actions are implemented.

2. Underlying Principles

The underlying *Principles Governing International Statistical Activities* upon which the DQAF is based were formulated by the Committee for the Coordination of Statistical Activities and endorsed by the chief statisticians/coordinators of statistical activities of UN agencies in 2005. They are as follows.

1. ***High quality international statistics, accessible for all, are a fundamental element of global information systems***

Good practices include:

- Having regular consultations with key users both inside and outside the relevant organisation to ascertain that their needs are met
- Periodic review of statistical programmes to ensure their relevance
- Compiling and disseminating international statistics based on impartiality
- Providing equal access to statistics for all users
- Ensuring free public accessibility of key statistics

2. ***To maintain the trust in international statistics, their production is to be impartial and strictly based on the highest professional standards***

Good practices include:

- Using strictly professional considerations for decisions on methodology, terminology and data presentation
- Developing and using professional codes of conduct
- Making a clear distinction, in statistical publications, between statistical and analytical comments on the one hand and policy prescriptive and advocacy comments on the other

3. ***The public has a right to be informed about the mandates for the statistical work of the organisations***

Good practices include:

- Making decisions about statistical work programmes publicly available
- Making documents for and reports of statistical meetings publicly available

4. ***Concepts, definitions, classifications, sources, methods and procedures employed in the production of international statistics are chosen to meet professional scientific standards and are made transparent for the users***

Good practices include:

- Aiming continuously to introduce methodological improvements and systems to manage and improve the quality and transparency of statistics

- Enhancing the professional level of staff by encouraging them to attend training courses, to do analytical work, to publish scientific papers and to participate in seminars and conferences.
- Documenting the concepts, definitions and classifications, as well as data collection and processing procedures used and the quality assessments carried out and making this information publicly accessible
- Documenting how data are collected, processed and disseminated, including information about editing mechanisms applied to country data
- Giving credit, in the dissemination of international statistics, to the original source and using agreed quotation standards when reusing statistics originally collected by others
- Making officially agreed standards publicly available

5. Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimise the reporting burden for data providers

Good practices include:

- Facilitating the provision of data by countries
- Working systematically on the improvement of the timeliness of international statistics
- Periodic review of statistical programmes to minimise the burden on data providers
- Sharing collected data with other organisations and collecting data jointly where appropriate
- Contributing to an integrated presentation of statistical programmes, including data collection plans, thereby making gaps or overlaps clearly visible
- Ensuring that national statistical offices and other national organisations for official statistics are duly involved and advocating that the Fundamental Principles of Official Statistics are applied when data are collected in countries

6. Individual data collected about natural persons and legal entities, or about small aggregates that are subject to national confidentiality rules, are to be kept strictly confidential and are to be used exclusively for statistical purposes or for purposes mandated by legislation

Good practices include:

- Putting measures in place to prevent the direct or indirect disclosure of data on persons, households, businesses and other individual respondents
- Developing a framework describing methods and procedures to provide sets of anonymous micro-data for further analysis by bona fide researchers, maintaining the requirements of confidentiality

7. Erroneous interpretation and misuse of statistics are to be immediately appropriately addressed

Good practices include:

- Responding to perceived erroneous interpretation and misuse of statistics
- Enhancing the use of statistics by developing educational material for important user groups

8. Standards for national and international statistics are to be developed on the basis of sound professional criteria, while also meeting the test of practical utility and feasibility

Good practices include:

- Systematically involving national statistical offices and other national organisations for official statistics in the development of
- international statistical programmes, including the development and promulgation of methods, standards and good practices
- Ensuring that decisions on such standards are free from conflicts of interest, and are perceived to be so
- Advising countries on implementation issues concerning international standards
- Monitoring the implementation of agreed standards

9. Coordination of international statistical programmes is essential to strengthen the quality, coherence and governance of international statistics, and avoiding duplication of work

- Good practices include:
- Designating one or more statistical units to implement statistical programmes, including one unit that coordinates the statistical work of the organisation and represents the organisation in international statistical meetings
- Participating in international statistical meetings and bilateral and multilateral consultations whenever necessary
- Working systematically towards agreements about common concepts, classifications, standards and methods
- Working systematically towards agreement on which series to consider as authoritative for each important set of statistics
- Coordinating technical cooperation activities with countries between donors and between different organisations in the national statistical system to avoid duplication of effort and to encourage complementarities and synergy
- to the improvement of statistics in the organisations and in countries

10. Bilateral and multilateral cooperation in statistics contribute to the professional growth of the statisticians involved and to the improvement of statistics in the organizations and in countries

Good practices include:

- Cooperating and sharing knowledge among international organisations and with countries and regions to further develop national and regional statistical systems
- Basing cooperation projects on user requirements, promoting full participation of the main stakeholders, taking account of local circumstances and stage of statistical development
- Empowering recipient national statistical systems and governments to take the lead
- Advocating the implementation of the Fundamental Principles of Official Statistics in countries
- Setting cooperation projects within a balanced overall strategic framework for national development of official statistics

3. Quality Dimensions

Introductory Remarks

It is generally agreed that, whilst *statistical product quality* can be summarized in line with the definition in the ISO 9000 Series for any product as *fitness for use*, there is a need to elaborate this definition in terms of the various quality aspects or *dimensions*. Many versions of quality dimensions have been proposed over the last 20 years, most of which contain essentially the same ideas and all of which include a significant expansion of the original narrow interpretation of quality as simply *accuracy*.

In developing the following set of dimensions particularly influential documents are:

- the *OECD Quality Framework and Guidelines* (the primary source);
- the generic *National Quality Assurance Framework (NQAF) Template* developed by UN Statistical Division and endorsed by the UN Statistical Commission; and
- the European Statistical System quality dimensions that were subsequently incorporated in the *European Statistics Code of Practice*.

The ITU quality dimensions are in two groups: those relating to *data quality* and those relating to *process quality*. The latter group is important as well designed and executed processes provide the basis for data quality.

Data Quality Dimensions

1. Relevance

The *relevance* of a data product is the degree to which the data serve to address the purposes for which they are sought by users. Relevance has three aspects: coverage of the required population (completeness); inclusion of the appropriate content: and use of appropriate concepts. Relevance is further characterised by the merit of the data uses in terms of the ITU mandate.

Typically a data product has multiple users and uses. Thus, measuring relevance requires the identification of user groups and their needs, whilst recognizing that these may change over time.

Relevance may be indirectly assessed by ascertaining whether there are processes in place to determine the views of users and the uses they make of the data.

Users of ITU data may be divided into two main groups;

- *internal users* - primarily analysts within ITU divisions; and
- *external users* - including other UN organisations, other international organisations, national governments, national statistical offices and other national organisations, academic institutions and the media.

Whilst internal users are the very important, it is essential that the content and format of published outputs be adapted to the full range of audiences.

2. Accuracy

The *accuracy* of a data product is the degree to which the data correctly estimate or describe the quantities or characteristics they are designed to measure. Accuracy refers to the closeness between the values provided in the product and the (unknown) true values. Accuracy has many attributes, and in practical terms there is no single overall measure of it. Typically, accuracy is

described in terms of the errors, or the potential significance of errors, introduced at various stages in the production process from initial acquisition of the data to dissemination of aggregates.

In the case of data from sample surveys, the major sources of error are coverage, sampling, non-response, response, processing, and seasonal adjustment. For data from censuses there are no sampling errors. For data from administrative sources, there are also no sampling errors, but there are additional problems due to mismatching of administrative concepts or classifications to statistical requirements.

An aspect of accuracy, commonly referred to as *reliability*, is the closeness of the initially released values to the subsequent values of data releases. In this context it useful to consider the sources of revision, which include (1) replacement of preliminary source data with later data, (2) replacement of projections with source data, (3) changes in definitions or estimating procedures, and (4) updating of the base year for constant-price estimates.

The accuracy of the data produced by ITU is largely determined by the accuracy of the data received from the contributing organisations. ITU activities can improve accuracy; for example, quality checks may detect errors in data provided by contributing organisations and lead to improvements in these data. Alternatively ITU activities can have an adverse effect, for example by introducing errors during the processing stages.

3. Credibility

The *credibility* of a data output refers to the confidence that users place in that product based primarily on their image of the data producer and the product, i.e., the *brand image*. It is based on the users' *perceptions* of accuracy as well as the actual accuracy.

Credibility is built over time. An important aspect is trust in the objectivity of the data. This implies that the data are perceived to be produced professionally in accordance with appropriate statistical standards, and that policies and practices are transparent. In particular, data are not manipulated, nor their release timed in response to political pressure.

Another aspect of credibility is trust in the integrity of the production process. To obtain complete coverage ITU may impute data for missing countries; to improve accuracy it may adjust data received. The extent to which this is well done and well understood affects credibility. Also, once agreement between ITU and an organisation has been reached on how data will be provided or imputed, the agreement should not be subsequently withdrawn in response to political pressure.

4. Coherence

The *coherence* of a data product reflects the degree to which it is logically connected and mutually consistent with other data products. Coherence implies that the same term should not be used without explanation for different concepts or data items; that different terms should not be used without explanation for the same concept or data item; and that variations in methodology that might affect data values should not be made without explanation.

Coherence in its loosest sense implies the data are "at least reconcilable." For example, if two data series purporting to cover the same phenomena differ, the differences in time of recording, valuation, and coverage should be identified so that the series can be reconciled.

Coherence has four important sub-dimensions.

- *Coherence within a dataset* implies that the elementary indicators are based on compatible concepts, definitions, and classifications and can be meaningfully combined. Incoherency within a dataset occurs, for example, when indicator values that should add up to a total do not.
- *Coherence across datasets* implies that the data are based on common concepts, definitions and classifications, or that any differences are explained and can be allowed for. An example of incoherency across datasets would be household ICT usage could not be reconciled with ICT supply. Unexplained inconsistencies across datasets can seriously reduce the interpretability and credibility of ITU statistics.
- *Coherence over time* implies that the data are based on common concepts, definitions, and methodology over time, or that any differences are explained and can be allowed for. Incoherence over time refers to breaks in a series resulting from changes in concepts, definitions, or methodology.
- *Coherence across countries* implies that, from country to country, the data are based on common concepts, definitions, classifications and methodology, or that any differences are explained and can be allowed for. Ensuring coherence across countries, commonly referred to as *harmonization*, is one of the major sources of value added by ITU.

Metadata plays a fundamental role in explaining possible changes in concepts or methodologies over time and across countries.

5. Timeliness and Punctuality

The *timeliness* of a data product is the length of time between its availability and the event or phenomenon it describes. Timeliness is assessed in terms of a time scale that depends upon the period for which the data are of value, i.e., are sufficiently timely to be acted upon. The concept applies equally to short-term or structural indicators, the only difference is the time scale.

Although ITU processes themselves can have an adverse effect, for the most part the timeliness of ITU data products is determined by the timeliness of the data it receives from the contributing organisations.

Punctuality implies the existence of and adherence to a *data product dissemination schedule.* A data product is punctual if it is disseminated in accordance with the schedule. In the case of data published externally the schedule may comprise a set of target release dates or may involve a commitment to release data within prescribed time period. (Here "release date" refers to the date on which the data are first made publicly available, by whatever medium, typically, but not inevitably the web site).

A dissemination schedule assists:

- internal users, by enhancing their capacity to plan their work based on target internal dissemination dates for data they require;
- external users, by improving their capacity to make timely use of ITU statistics;

There may be occasions when ITU simply cannot adhere to the dissemination schedule due to the late acquisition of data from input sources. In such circumstances advance warning regarding the delay in dissemination should be communicated to users.

Although timeliness and punctuality are different concepts they are grouped together in a single quality dimension for two reasons, first because their separate achievements are heavily inter-related in practice, and second to be in line with international standards and practices.

6. Accessibility

The *accessibility* of a data product reflects how readily the data can be discovered, located and accessed from within ITU data holdings. It includes the suitability of the formats in which the data are available, the media of dissemination, the availability of metadata and user support services, and, in the event that there is a charge, the affordability of the data to users.

From the perspective of data availability, ITU users are divided into two very distinct groups: internal users; and external users. Typically, because of the differences in access methods, internal users can access data earlier and in more detail than external users. Thus these two groups may have quite different perceptions of accessibility.

The range of different external users leads to the need for multiple dissemination formats and selective presentation of metadata. A publication policy should be articulated and made publicly known.

7. Interpretability

The *interpretability* (sometimes called *clarity*) of a data product reflects the ease with which users can understand and properly use the data. The degree of interpretability is largely determined by the adequacy of the definitions of concepts, target populations, indicators and other terminology describing the data, and its limitations.

If there are several dissemination mechanisms they should be harmonised in order to avoid confusing users.

Coping with the needs of the broad range of external users leads to the use of metadata presentation in layers of increasing detail. The content and format of published products should be adapted to the different target groups.

Balancing Data Quality Dimensions

The data quality dimensions are not mutually exclusive in the sense that there are relationships between the factors that contribute to them. Factors leading to improvements with respect to one dimension may result in deterioration with respect to another. Thus, in designing a data collection and products, it is often necessary to trade-off quality in one dimension with quality in another. The most significant trade-offs to consider are as follows.

- *Accuracy and timeliness.* This is probably the most frequently occurring and important quality trade-off. Improvement in timeliness can be obtained by terminating data acquisition earlier and compiling products based on a smaller number of countries and/or reduced editing. However, as this reduces accuracy, there needs to be a trade-off. For major products a compromise is to disseminate a preliminary version of the data product based on partial acquisition and then one or two revised products based on successively more acquisition and editing. The size of the revisions between preliminary and revised products is an indicator of degree of accuracy that is being sacrificed in order to produce the increased timeliness.
- *Relevance and accuracy*. Relevance can be increased by acquiring more data items, but accuracy may be diminished because the additional data are less reliable. Conversely elimination of inaccurate data items will increase accuracy but reduce relevance.

- *Relevance and timeliness*. Timeliness may be improved by reducing the number of data items collected or by replacing those that are difficult to collect by ones that are easier. This will have a negative effect on relevance.
- *Relevance and coherence*. Improvements in relevance, for example by redefining the indicators for which data are collected, or moving to a later version of a classification, will reduce comparability over time, perhaps to the point of requiring a series break. Conversely, the desire to retain comparability over time may inhibit changes in content required to improve relevance.
- *Accuracy and coherence*. Improved methods may increase accuracy but reduce coherence by introducing changes in data that are attributable to changes in methods not in what is being measured. Conversely, the desire to retain coherence may inhibit the changes required to improve accuracy.

Process Quality Dimensions

8. Sound Methodology

Sound methodology refers to the use of international standards and best practices through all stages of a data collection from identification of requirements, through design, data collection, processing, analysis, dissemination, archiving and evaluation. Application of standards and best practices not only engenders ITU process and product quality, it fosters comparability across organisations and countries.

Sound methodology includes both theory and its application in the sense of ensuring that, not only are procedures well designed, but also they are well implemented and documented, and that staff are well informed and trained.

9. Sound Systems

Sound systems refers to the use of international standards and best practices in systems development, including liaising with systems developers in other statistical organisations and making optimum use of off-the-shelf or shared statistical products where available.

Sound systems also includes both theory and its application in the sense of ensuring that systems are well designed, developed, implemented and documented, and that staff are well trained in their use.

10. Cost-Efficiency

The *cost-efficiency* with which data products are produced is a measure of the costs incurred and resources expended relative to the benefits of the products. The aim is to produce a specified set of products of specified quality at minimum cost.

Efficiency can affect all dimensions of product quality in the sense that, if a product can be produced more efficiently with the same quality, then the resources released can be used to improve any dimension of the quality of that product or other products, or to create new products.

Two types of costs are incurred:

- the costs to ITU of acquiring, processing and disseminating the data;

- the costs incurred by the NSOs and NRAs from which the data are acquired. These costs depend significantly on whether or not these national organisations collect the data for their own purposes. If they do, then the costs are essentially those of repackaging and transmitting data already collected. Otherwise they are the full costs of data collection.

ITU never collects data directly from basic units (operators, households, and individuals) to which the data refer and that provide the original elementary data. Thus their costs in responding are of only indirect concern.

ITU acquires ICT data for EU member countries from Eurostat. Eurostat is collecting this information in pursuance of an EU regulation and thus little or no additional burden is imposed by ITU.

4. Quality Guidelines

Introductory Remarks

The guidelines are presented in two broad groups: those applying to any *individual statistical production process* within ITU; and those applying to the ITU *statistical infrastructure.*

For the purposes of presenting guidelines for a statistical production process, the various activities associated with the collection are subdivided into eight subgroups corresponding to the *phases* defined in the *Generic Statistical Business Process Model (GSBPM),* published by UNECE, which is the international standard. The groups have been slightly expanded to include certain closely related statistical infrastructure activities, namely managing *provider relations, managing user relations and assuring quality*

The remaining activities associated with the statistical infrastructure are presented in three subgroups, namely, *assure methodology impartiality, objectivity and transparency; manage metadata,* and *manage human, financial and technological resources.*

Even though not explicitly referenced, staff training is part of every subgroup.

In principle, for each subgroup of statistical activities thus defined, the guidelines should include the following:

1. *Scope* – a short description of the statistical activities to which the guidelines refer;
2. *General guidelines* – statements of best practice reflecting the aims of the guidelines in general terms;
3. *Detailed guidelines* – covering all aspects of quality and performance to be addressed;
4. *Monitoring mechanisms* – the methods by which adherence to the guidelines might be monitored, including quality and performance indicators and quality assessments; and
5. *Reference documentation* – documents that elaborate the guidelines and/or that were instrumental in their formulation.

This first version of the DQAF contains only Scope and General Guidelines.

The General Guidelines may also be used as a self-assessment checklist. The degree to which current activities correspond to each statement of best practice can be assessed on a five point scale:

- +2: completely

- +1: for the most part
- 0: not applicable
- -1: to a limited extent
- -2: not at all

The scores can be aggregated across each activity group, and over all groups, to give summary quality scores that can be monitored over time, or compared with other statistical production processes and products assessed using the same checklist.

1. Specify Needs and Manage Users

Scope

This group includes all activities associated with engaging users to identify their statistical needs, proposing strategic options for meeting these needs, and, if required, preparing a business case for changes to the production process and products. It includes: maintaining a knowledge of, and good relationships, with users, and examining their statistical needs in detail; checking the extent to which current statistical production process and other data sources meet these needs; defining the general content of new or changed statistical products to meet the needs; and preparing a business case where needed to secure approval and resources to re-engineer the production process and/or produce new or changed statistics

General Guidelines

ITU has:

- recently determined the needs, or changes in needs, for the statistical data products it produces;
- categorised the users and potential users by type of use and data needs;
- reviewed user requests, queries and comments with a view to identifying new or changing user needs.
- identified the most important, say 20, individual users and discussed their needs and specified in detail the corresponding data requirements;
- established memoranda of understanding, service level agreements or equivalent for provision of data to key users;
- conducted user satisfaction surveys on a regular basis;
- identified the relevant populations and indicators for which data are required and the appropriate definitions and classifications;
- evaluated data currently available from other sources and determined the extent to which the data requirements can be met with data from these other sources;
- prepared the business case for development, or substantial revision, of the statistical process and products;
- consulted with key users on proposed changes to statistical products;
- shared information about new or revised statistical products with ITU divisions thereby maximizing the possibility of data coherence and minimizing the risk of duplication of effort and waste of resources;

- provided opportunities to interested staff in different divisions to contribute to the development or redevelopment of the data production process;
- maximised use of existing data sources before embarking on data additional collection;
- facilitated sharing of ICT related data between divisions within ITU;
- ensured the coherence of all ICT related data acquired and produced by ITU;

2. Design Statistical Production Process and Statistical Infrastructure

Scope

The statistical activities in this group follows from, and build on, the results of specifying user needs. They include the research and design work needed to define or redefine, as needed, the statistical concepts and indicators, and the data collection, processing, storage and dissemination procedures required to produce the envisaged statistical products. They include specification of the metadata that are inputs to and outputs from to the production process, and general design of the systems and workflows that support and enable the efficient conduct of the procedures.

General Guidelines

ITU has:

- defined and justified the scope and content, i.e., countries to be included, indicators and classifications;
- identified and used the most appropriate concepts, definitions, and classifications, where possible taking advantage of those already developed and used by the international statistical community;
- identified and reviewed all possible available data sources and selected the most appropriate;
- designed efficient and effective data collection systems and procedures, including making formal agreements with providers where required;
- designed efficient, effective and integrated data processing procedures and systems for coding, verifying, imputing, estimating, integrating, and disseminating data products, using internationally accepted methods to the fullest extent possible;
- designed statistical products, including tables, datasets, databases and analyses in accordance with the specified data requirements;
- identified and included all the metadata required to support data collection and production and to inform users;
- involved experts in data collection and processing, and in ICT statistics, wherever available;

3. Build Statistical Procedures and Systems

Scope

The statistical activities in this group cover building procedures and systems to support the production process and infrastructure and testing them prior to production. They include building the data collection tools, data repositories, processing tools, management control functions, and metadata management tools. They include configuring workflows to handle data

transformations from acquisition to archiving. They include producing documentation and training production staff, providers, and users in use of the systems and procedures.

General Guidelines

ITU has:

- built efficient data collection tools;
- built appropriate and efficient data storage mechanisms, for example using databases rather than worksheets for data storage;
- built efficient and effective processing tools and metadata management tools;
- configured smooth workflows in which all activities within the entire production process fit together efficiently without gaps or redundancies;
- documented all systems and procedures;
- trained production staff, providers and users in use of the systems;
- recently specified and met deadlines for introduction of new statistical products and/or reengineering of statistical processes;
- and if so, tested all new procedures and systems before putting them into production.

4. Collect Data and Manage Provider Relations

Scope

The statistical activities in this group refer to the actual acquisition of data, using the various sources and collection modes specified in the design phase, and storing the data acquired securely in an appropriate repository. They include implementing procedures and systems for data collection from NSOs, NRAs and databases of other international or commercial organisations, including follow-up in the event of non-response or dubious data values. They include liaising with providers, making them aware of the reasons for and specifics of the data required, and responding to their comments, queries and complaints. They include ensuring that the data and associated metadata are loaded into a suitable repository.

General Guidelines

ITU has:

- ensured that NRAs and NSOs are totally familiar with the reasons for data collection and the precise meanings of the indicators to be provided?
- implemented efficient and effective procedures for collecting data and associated metadata from these providers?
- made provision for responding to comments, queries and complaints from these providers?
- monitored and followed up non-responses and partial responses and thus ensured that as much of the required data and metadata are collected?
- where appropriate, implemented procedures to obtain data and corresponding metadata from the databases of other international or commercial organisations?
- ensured the risk of errors in data and metadata acquisition is minimised?

- minimised the reporting burden on organisations providing data and metadata?
- minimised ITU resources spent in data and metadata acquisition.
- ensured that all information concerning data and metadata flows between any organization within a country and ITU is reported to the country NSO (in view of its role as statistical coordinator for the country);
- implemented procedures that ensure data and metadata are loaded into a suitable repository;

5. Process Data

Scope

The statistical activities in this group refer to the verification, editing, harmonisation and imputation of incoming data and their preparation for analysis. They include integrating data from various sources and classifying and coding them where needed; applying checks that identify missing, invalid or inconsistent data or metadata; imputing missing values for which no data have been received or for which data are inadequate or in error; applying adjustments to harmonise data across countries; deriving values for indicators for which data are directly acquired, compiling totals, averages and ratios for regions, and measures of dispersion; and storing data and metadata in databases from which data can be readily extracted for analysis purposes, statistical products can be compiled, and users internal to ITU can make extracts.

General Guidelines

ITU has:

- implemented efficient and effective procedures and systems for data collection;
- implemented efficient and effective procedures and systems for data, verification, editing, harmonisation, including a full range of validity checks and edits;
- implemented efficient and effective procedures and systems for derivation and imputation;
- implemented efficient and effective procedures and systems for aggregation, including production of totals, averages and ratios for regions, and measures of dispersion;
- implemented efficient and effective procedures and systems for storage of data and metadata in an internal database from which statistical products can readily be compiled and internal users can make extracts;

ITU does:

- minimize the human resources required for data and metadata processing by good workflow management;
- collect, review and analyse all operational metadata, including non-response rates, indicator non-response rates, indicator imputation rates;
- facilitate internal user access to the data;
- facilitate internal user understanding of the data, and their limitations, through provision of metadata describing procedures and operations;
- facilitate integration of the data with data from other production processes and checked their coherence;

6. Analyse Data

Scope

The statistical activities in this group refer to the analysis required for verification of the statistical products and their preparation for dissemination. They include preparing draft statistical products, including associated metadata and quality indicators and checking that they will support the analyses for which they were designed. They include undertaking the analyses required to explain the data. They include scrutinizing, analysing and explaining the data in relation to expectations and identifying divergences from expectations. They include finalising the statistical products including interpretation notes, briefings, measures of uncertainty and other relevant metadata. (In the event that statistical products include data for individual operators or very small groups of operators, the activities also include checking that the data did not breach any relevant confidentiality rules.)

General Guidelines

ITU does:

- prepare and review draft data products and scrutinize, analyse and explain the data in relation to expectations;
- compare the data with data for previous reference periods, and confront the data with any related data from other sources;
- view the data from all other perspectives and ensured there is an in-depth understanding of the data content before dissemination;
- produce and analyse quality indicators;
- discuss the results with internal experts and otherwise check that statistical products and associated metadata are fit for purpose before external dissemination;

7. Disseminate Statistical Products

Scope

The statistical activities in this group refer to the dissemination of the statistical products to users within ITU and to external users. They include formulating and applying a dissemination strategy, including a release schedule and pricing policy; reformatting the data and metadata as required and loading them into publicly accessible databases, and preparing and disseminating printed publications in accordance with ITU publishing and presentation guidelines; notifying users of the release of the statistical products, giving briefings for key users and user groups; promoting the products to ensure that they reach the widest possible range of users; and managing communications with internal and external users, including ensuring that user queries are recorded and that responses are provided

General Guidelines

ITU has:

- implemented efficient and effective procedures and systems for data dissemination;

ITU does:

- notify internal users and enable their access to statistical products at the earliest possible opportunity;
- prepare statistical products in accordance with relevant ITU publishing and presentation guidelines;
- maximize internal and external interpretability of the data products by accompanying them with appropriate metadata;
- ensure that products are timely and punctual by maintaining and adhering to detailed activity and release schedules;
- disseminate statistical products externally, including preparing printed publications and making databases accessible via the web site;
- give briefings to key users, including senior officials of international and national organisations, large ICT service providers, members of think tanks, academia and the media;
- promote externally products via wikis and blogs and social media to ensure that they reach the widest possible audience, for example;
- manage communications with internal and external users by ensuring that user queries are recorded, that responses are provided;

8. Evaluate and Assure Quality

Scope

The statistical activities in this phase focus on evaluation (assessment) of the statistical process products and infrastructure within the context of data quality assurance framework. They include developing and implementing the framework, including quality principles, dimensions, guidelines and evaluation procedures.

Depending upon the scale of the evaluation, activities may include: defining the evaluation objectives and procedures; establishing the evaluation team; assembling the relevant documentation; analysing the documentation; discussing the statistical process and products with the staff responsible for them and with key users and providers; identifying quality and performance issues; and making recommendations for improving quality and performance.

The evaluation may refer to one specific repetition (cycle, instance) of the process and product, or to a particular set of repetitions, for example the process and all the products during the course of the previous five years. The evaluation may be lightweight or detailed, ranging from self-assessment of the values of key quality and performance indicators for a single repetition to a comprehensive external assessment. In its most detailed form, evaluation involves a complete review of the products relative to the original data requirements and of the production process relative to the original design.

General Guidelines

ITU has:

- reviewed quality concepts and prepared a data quality assurance framework, including quality principles, dimensions, guidelines and assessment checklist;
- defined evaluation procedures and implemented an evaluation programme;
- recently conducted a self-assessment of its production process and/or products;

- recently commissioned an external assessment of its production process and/or products;

In the event that an evaluation has been recently conducted, ITU did:

- set the objectives and scale of the evaluation;
- provide the evaluation team with comprehensive documentation;
- review and discuss the evaluation results;
- based on the results, implement quality and performance improvements;
- based on the results, prepare a business case for quality and performance improvements and submit to senior management for support and resources;

9. Assuring Sound Methodology, Impartiality, Objectivity and Transparency

Scope

ITU needs to be seen to collect, produce and disseminate statistics in a manner that is based on sound methods and that is professional, transparent, impartial and objective.

General Guidelines

ITU has:

- made known the commitment to follow professional standards in collecting, producing and disseminating statistics that are impartial and objective, including by publicizing its data quality assurance framework;
- developed and made public a data dissemination policy;
- developed and made public a release calendar in which dissemination dates and times are pre-announced and ensure that any deviations from the calendar are announced and justified to the users;

ITU does:

- base the recruitment and promotion of statistical staff on aptitude and expertise in statistics;
- ensure that statistics are produced on an objective basis, determined only by statistical considerations;
- select sources, concepts, methods, processes and dissemination practices on the basis of statistical considerations, and national and international principles and best practices;
- explain major changes in the methodologies and data revisions to users;
- ensure that statistical releases are clearly distinguished from political/policy statements and issued separately from them;
- ensure that statistical releases and statements made at press conferences are objective and non-partisan;
- correct errors in data outputs as soon as possible after they are detected and inform users;

10. Manage Metadata

Scope

In broad terms metadata are *data about data*. In the specific context of quality assessment, metadata are data about every aspect of the IDS statistical production process and its products. They are divided into four broad groups:

- *Data-related metadata* are those metadata that directly describe the data as they are input and output during the various phases constituting the production process. Examples are descriptions of rows and columns in output tables, and descriptions of the contents of a database
- *Definitional metadata* are the metadata that describe the concepts and definitions used in the production process. Examples are definitions of indicators, description of target and actual populations, question wordings in questionnaires.
- *Procedural metadata* are those metadata that describe the particular procedures that make up the phases of the production process. An example of procedural metadata is the specification of the automated verification rules applied during data entry. Procedural metadata are further divided into two types: *active* procedural metadata drive a procedure in the sense that the procedure cannot commence without them; and *passive* procedural metadata simply document the procedure
- *Operational metadata* are metadata that describe the inputs and outputs of a procedure, other than the actual data. Operational metadata are also further divided into two types. *Input operational metadata* comprise metadata that *enable and control* the execution of each particular instance of a procedure. *Output operational metadata* comprise metadata that *result from the execution* of a particular instance of a particular process. They include information to be passed to a later activity. They also include *process metrics*, also called *paradata*, that are generated during the various production phases. Such metadata are the source of quality and performance and quality indicators.

Activities associated with managing metadata include identifying and classifying the various types of metadata of interest; determining how they are to be managed, i.e., obtained, recorded, accessed and used; building appropriate procedures and systems to do so; and analysing them in order to better understand and improve the production process and products.

General Guidelines

ITU has:

- identified and fully documented the various types of metadata, and the needs for and uses of them;
- designed, built and operated a comprehensive metadata management infrastructure;
- appointed a single, authoritative registration authority for each metadata type;
- registered metadata using a registration process that is well documented so there is clear identification of ownership;
- ensured that metadata are active to the greatest extent possible, thereby ensuring they are accurate and up-to-date, and paving the way for automation;

- ensured that passive metadata are recorded at the time they are created, preferably automatically as a by-product of the processes that generate them;
- ensured that there is a single copy of each metadata value, which is entered once and can be accessed or superseded, but not overwritten, earlier values being retained to allow historical access;
- constructed different views of metadata corresponding to the differing needs of the various users;
- reused metadata wherever possible rather than recreating them.

11. Manage Human, Financial and Technological Resources

Scope

Managing resources involves ensuring that the financial, human, and technological resources available for statistical work are adequate both in magnitude and quality, and are sufficient to meet the needs associated with the development, collection, production and dissemination of statistics.

General Guidelines

ITU does:

- make maximum use of ICT in supporting the statistical process;
- measure and analyse the work effort involved in each phase of the production process;
- design and conduct statistical activities in each phase such a way as to promote efficiency;
- ensure user feedback is taken into account in determining priorities when resources are limited;
- ensure that the financial and human resources available match the statistical activities envisaged;
- review resource allocations on a regular basis.

5 Quality Assessment and Improvement Programme

Introductory Remarks

Quality has to be managed through a comprehensive quality assessment and improvement programme that draws on and puts into action the quality principles, dimensions and guidelines described above. The programme comprises three types of quality assessment:

- monitoring quality and performance indicators for each repetition of each statistical production process;
- annual or biennial quality self-assessment of each production process or component of the statistical infrastructure;
- external quality review of a particular process or infrastructure component on an occasional basis.

These types of assessment differ from one another in their aims, the amount of detail and effort involved, the frequency with which they are conducted and the sort of recommendations they

may generate. In all cases the goal is to improve quality. So there has also to be a process for ensuring that all recommendations arising from an assessment are implemented where this can be done with existing resources or are advanced to senior management for consideration if addition resources are required.

Monitoring Quality and Performance Indicators

The objectives of identifying and monitoring quality and performance indicators (QPIs) are to quickly check ongoing operations, to monitor performance with respect to target objectives, and to identify sources of operational errors and correct them.

QPIs monitor statistical operations in terms of quality (i.e., effectiveness) and performance (i.e., efficiency). They may be divided into two groups:

- *product QPIs* - monitoring output indicators and analyses;
- *process QPIs* - monitoring all phases of statistical processes and infrastructure;

QPIs for the DQAF have to be developed by the staff with intimate knowledge of the statistical activities to be monitored. They must be very carefully chosen. Too few QPIs, or the absence of QPIs for key procedures or outputs, result in ineffective monitoring. Too many QPIs, or ill-chosen ones, overload the production procedures and are a waste of resources.

The procedures involved in development and use of QPIs are:

- define a preliminary set of QPIs;
- designate selected QPIs as being *key* and set targets for each of these;
- analyse the values of process and product QPIs for each repetition of each statistical process;
- take immediate action to address the *operational* problems thereby identified; and
- document *structural* problems, i.e., problems that cannot be solved at operational level, and provide them as input to the next quality self-assessment, and, if serious, let them trigger a quality self-assessment.

Quality Self-Assessment of a Statistical Production Process and Its Products

A self-assessment of quality is conducted by the staff responsible for the statistical process on an annual basis. Its objectives are to help the staff responsible to develop an impression of the quality of the process and products, and hence to identify structural weaknesses and to propose quality improvements. It involves:

- Assembling documentation about the process and its products;
- Convening one or more meetings with the staff responsible for all aspects of processing, and at these meetings reviewing the documentation, completing the relevant sections of the checklist, and identifying process and product weaknesses and potential improvements;
- Convening one or more meetings with the principal users and at these meetings reviewing the products, completing the relevant sections of the checklist and identifying product weaknesses and potential improvements; and
- Taking action on any improvements that can be implemented with existing resources and documenting improvements that would require additional resources and/or support from other areas.
- Presenting a summary of the results to senior management.

Self-assessment is facilitated by use of a self-assessment checklist. The checklist should cover all aspects of the production process from identification to dissemination each viewed in terms of the quality dimensions. As previously noted, the quality guidelines

External Quality Assessment

External quality assessment is appropriate on a regular but infrequent basis or if concerns about quality of products or processes reach a high level. The assessment objectives are to provide Division Head and ITU senior management and with an objective view of the quality of the statistical production process, and hence to identify any structural weaknesses, to propose quality improvements to address them, and to indicate the resource implications.

An external quality assessment involves the following steps:

- Defining the terms of reference for the assessment;
- Identification of the assessment team, involving an external expert, with the manager of the data collection as a resource person;
- Obtaining documentation and the results of recently completed self-assessments;
- Convening meetings with relevant ITU staff to further elaborate the problem areas and improvements required to address them;
- Convening meetings with the principal users and further investigating the problem areas as reflected in product weaknesses;
- Reporting the results of the assessment to ITU management in accordance with the terms of reference.