**Session 1: Introduction: Background and current state of the art of quality assurance frameworks in international organizations and at the national level (principles, codes of practices, ongoing initiatives, existing handbooks, guidelines and other tools)**

# IMF's Data Quality Assessment Framework[1]

## I. INTRODUCTION

This paper describes the Data Quality Assessment Framework (DQAF), how it was developed by the International Monetary Fund (IMF) and its structure (http://dsbb.imf.org/Applications/web/getpage/?pagename=dqrshome). It also illustrates the application of the DQAF in several important statistical efforts including the Fund's technical assistance and data quality assessments, as well other international efforts to support countries' statistical practices. The paper's structure follows chronologically the inception and further development of the Fund's DQAF, keeping the focus on the DQAF as a powerful tool for assessing data quality.

The first section provides background information on the Fund's DQAF. The second section of the paper covers the development of the DQAF and the consultation process. The third section deals with the structure of the DQAF. The fourth section covers the different uses of the DQAF by the IMF and other users to assess the quality of data, and the last section contains a conclusion.

## II. WHAT IS THE DQAF

The generic DQAF July 2003 serves as an umbrella for seven dataset-specific frameworks. The DQAF July 2003 was introduced at the Fifth Review of the Fund's Data Standards Initiatives. The DQAF is at the center of the IMF Data Quality Program (DQP). The DQP is a

---

[1] This paper was drafted by a team in the IMF's Statistics Department including Ethan Weisman, Zdravko Balyozov, and Louis Venter.

set of well-integrated initiatives, which comprises applications of the DQAF, projects to support and promote good statistical practices identified in the DQAF, and the maintenance and development of dataset-specific DQAFs.

The DQAF is the organizing model of the Fund's activity regarding the data modules of the Report on the Observance of Standards and Codes (Data ROSC). To date, over 100 Data ROSCs have been published. They cover a broad mix of industrial and developing countries. In addition to its use in Data ROSCs, the DQAF has been applied in the Fund's statistical TA program and in the statistical capacity building indicators developed under the auspices of the Partnership in Statistics for Development in the 21st Century (PARIS21). The DQAF is a framework that allows users and compilers to make their own data quality assessments. The generic framework brings together the internationally accepted core principles, standards, or practices for official statistics and serves as the umbrella under which the dataset-specific quality assessment frameworks were developed. The DQAF identifies quality-related features of governance of statistical systems, statistical processes, and statistical products. The DQAF is rooted in the UN Fundamental Principles of Official Statistics and grew out of the Special Data Dissemination Standard (SDDS) and General Data Dissemination System (GDDS), the IMF's initiatives on data dissemination.

## III. DEVELOPMENT OF THE DQAF

In a series of discussions, beginning in 1995, that examined data provision to the Fund, the IMF's Executive Board noted that it was imperative for the IMF, as well as for member countries, to improve the quality of the data used in policy analysis. This issue, of data quality, was again specifically addressed by the IMF's Executive Board on the occasion of the Third Review of the Fund's Data Standards Initiatives in March 2000, and again in June of that year during the discussion on Data Provision to the Fund for Surveillance Purposes.

In 1997 the Statistics Department (STA) of the IMF started to work on an approach to assessing data quality, with the development of a framework that was based on five broad areas that were considered relevant to an assessment of data quality across a wide range of uses and users. Building on this earlier work, STA developed the DQAF. The DQAF is a methodology for assessing data quality that brings together best practices and internationally accepted concepts and definitions in statistics, including those of the United Nations Fundamental Principles of Official Statistics and the SDDS/GDDS. The DQAF is a product of an intensive, iterative process of consultation with national statisticians, regional and international organizations, IMF staff, and data users outside the IMF, partially in response to the call for a substantial data quality assessment tool in the aftermath of the economic crises during the late 1990s.

In the development of the DQAF, two main areas of work were pursued to:

- Clarify the meaning and promoting an understanding of data quality in the community of data users and compilers; and

- Provide a structure and a common language for data quality that could be distilled into an assessment framework.

On the first point, the development of the DQAF took into account the consensus view that quality is a much wider, multidimensional concept than just accuracy. To promote a common understanding of data quality, STA established a Data Quality Reference Site on the Dissemination Standards Bulletin Board (DSBB)[2].

On the second point, the issue was approached by defining the characteristics that a data quality assessment framework needs to embody, such as:

- Comprehensive in coverage of the dimensions of quality, and of elements and indicators that might represent quality;

- Balanced between the rigor desired by an expert and the bird's-eye view desired by a general data user;

- Structured but flexible enough to be applicable across a broad range of stages of statistical development;

- Structured but flexible enough to be applicable (at least) to the major macroeconomic datasets;

- Designed to lead to transparent results; and

-  Arrived at by drawing on best practices of national statisticians.

After extensive consultations, the July 2001 DQAF that emerged brought together the internationally accepted core principles, standards, or practices for official statistics and provides structure and a common language for the assessment of data quality. The DQAF facilitates a comprehensive view of data quality, one that recognizes interrelations, including tradeoffs, among elements of quality and allows emphases to vary across data categories and uses/users. It facilitates dialogue with national statistical agencies and country authorities, as well as a more homogenous approach to assessing data quality by data compilers, data users, and Fund staff. The DQAF provides a structure for assessing existing practices against best practices, including internationally accepted methodologies.

At the time of the Fifth Review of the Fund's Data Standards Initiatives in July 2003 the IMF's Executive Board approved the refinement of the July 2001 DQAF to reflect experience and international statistical developments, which include the following:

- Good statistical practices identified through, the testing of the DQAF in the field and developments in the SDDS and the GDDS;

---

[2] http://dsbb.imf.org/Applications/web/getpage/?pagename=dqrshome

- Methodological improvements made at both the international level (e.g., the *Government Finance Statistics Manual 2001*) and the regional level (e.g., European Union guidelines); and

- Work with other international organizations on harmonizing approaches to data quality.

Refinements have been made to fill some gaps at the most detailed level and to address overlapping coverage in a few areas. Most refinements entailed introducing greater precision in the description of good statistical practices (e.g., separate assessment of the adequacy of staff resources, facilities, and computer resources). Drawing from the range of countries' experiences, greater attention was given to the effectiveness and efficiency of the management of statistical processes.

The generic DQAF July 2003 serves as an umbrella for eight dataset-specific frameworks, namely DQAF July 2003 DQAFs covering separately:

1. National accounts statistics;

2. Consumer price index;

3. Producer price index;

4. Government finance statistics;

5. Monetary statistics;

6. Balance of payments statistics;

7. External debt statistics; and

8. Household income in a poverty context[3]

## IV. HOW IS THE DQAF STRUCTURED?

The generic framework, as shown in Box 1, follows a cascading structure that flows from five main dimensions of data quality and a set of prerequisites for the assessment of data quality that have been identified as critical constituents of data quality. The DQAF is organized in a cascading structure that progresses from the abstract/general to the more concrete/specific details. The DQAF comprehensively covers the various quality aspects of data collection, processing, and dissemination.

---

[3] This dataset-specific DQAF was prepared in collaboration with the World Bank.

The first level covers the **prerequisites** of quality and five **dimensions** of quality: assurances of integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility. For each of these prerequisites and five dimensions, there are **elements** (two-digit level) and **indicators** (three-digit level)[4] (Box 2).

At the next level, **focal issues** that are specific to the compilation of a particular dataset (e.g. national accounts statistics) are addressed. Below each focal issue, **key points** identify quality features that may be considered in addressing the focal issues. The key points are meant to be suggestive, not exhaustive. Box 3 provides a view of the cascading structure employed in the Framework.

---

### Box 1.   The Data Quality Assessment Framework

The DQAF covers five dimensions of quality and a set of prerequisites for the assessment of data quality. The coverage of these dimensions recognizes that data quality encompasses characteristics related to the institution or system behind the production of the data as well as characteristics of the individual data product. Within this framework, each dimension comprises a number of elements, which are in turn associated with a set of desirable practices. The following are the statistical practices that are associated with each dimension:

**Prerequisites of quality**—the environment is supportive of statistics; resources are commensurate with needs of statistical programs; and quality is a cornerstone of statistical work.

**Integrity**—statistical policies and practices are guided by professional principles; statistical policies and practices are transparent; and policies and practices are guided by ethical standards.

**Methodological soundness**—concepts and definitions used are in accord with internationally accepted statistical frameworks; the scope is in accord with internationally accepted standards, guidelines, or good practices; classification and sectorization systems are in accord with internationally accepted standards, guidelines, or good practices; and flows and stocks are valued and recorded according to internationally accepted standards, guidelines, or good practices.

**Accuracy and reliability**—source data available provide an adequate basis to compile statistics; statistical techniques employed conform with sound statistical procedures; source data are regularly assessed and validated; intermediate results and statistical outputs are regularly assessed and validated; and revisions, as a gauge of reliability, are tracked and mined for the information they may provide.

---

[4] The first three levels are common with other data-set specific DQAFs that have been developed to assess datasets. This design was implemented to ensure a common and systematic assessment across datasets.

**Serviceability**—statistics cover relevant information on the subject field; timeliness and periodicity follow internationally accepted dissemination standards; statistics are consistent within the dataset, over time, and with other major data sets; and data revisions follow a regular and publicized procedure.

**Accessibility**—statistics are presented in a clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis; up-to-date and pertinent metadata are made available; and prompt and knowledgeable support service is available.
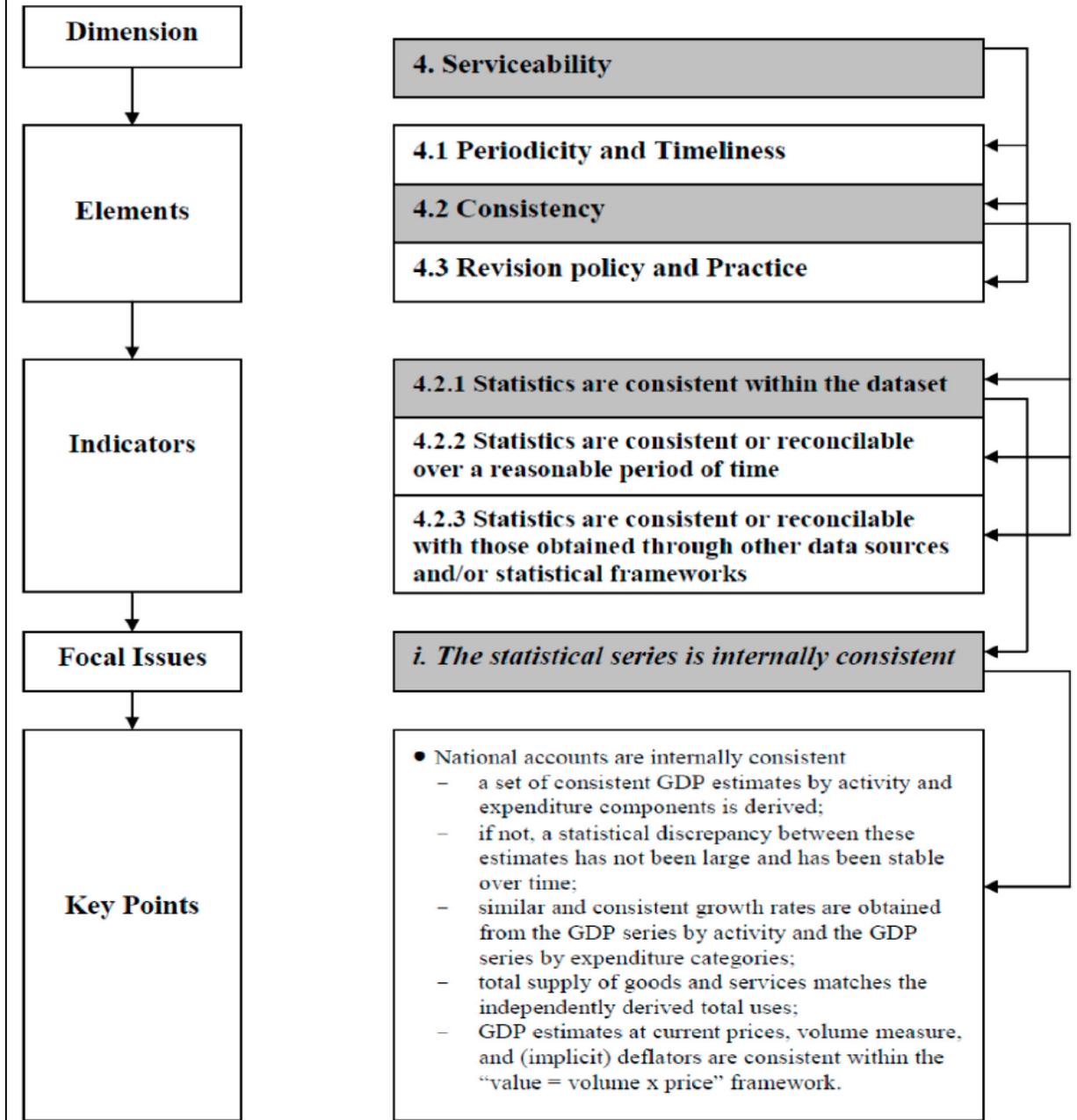
**Box 2.  Content of the Framework**

The elements and indicators within their respective dimensions are described below.

**0.**      **Prerequisites of quality**: Although not itself a dimension of quality, this group of "pointers to quality" includes elements and indicators that have an overarching role as prerequisites, or institutional preconditions, for quality of statistics. Note that the focus is on the agency, such as a national statistical office, central bank, or a ministry/department. These prerequisites cover the following elements:
    0.1      legal and institutional environment,
    0.2      resources available for the statistical program,
    0.3      relevance, and
    0.4      other quality management.

**1.**      **Assurances of integrity**: This dimension relates to the adherence to the principle of objectivity in the collection, compilation, and dissemination of statistics. The dimension encompasses institutional arrangements that ensure professionalism in statistical policies and practices, transparency, and ethical standards. The three elements for this dimension of quality are the following:
    1.1      professionalism,
    1.2      transparency, and
    1.3      ethical standards.

**2.**      **Methodological soundness**: This dimension covers the idea that the methodological basis for the production of statistics should be sound and that this can be attained by following internationally accepted standards, guidelines, or good practices. This dimension is necessarily dataset-specific, reflecting different methodologies for different datasets. This dimension has four elements, namely:
    2.1      concepts and definitions,
    2.2      scope,
    2.3      classification/sectorization, and
    2.4      basis for recording.

**3.**      **Accuracy and reliability**: This dimension covers the idea that statistical outputs sufficiently portray the reality of the economy. This dimension is also data specific, reflecting the sources used and their processing. The five elements of this dimension cover the following:
    3.1      source data,
    3.2      assessment of source data,
    3.3      statistical techniques,
    3.4      assessment and validation of intermediate data and statistical outputs, and
    3.5      revision studies.

**4.**      **Serviceability**: This dimension relates to the need that statistics are disseminated with an appropriate periodicity in a timely fashion, are consistent internally and with other major datasets, and follow a regular revision policy. The three elements for this dimension are as follows:
    4.1      periodicity and timeliness,
    4.2      consistency, and
    4.3      revision policy and practice.

**5.**      **Accessibility**: This dimension relates to the need for data and metadata to be presented in a clear and understandable manner on an easily available and impartial basis, that metadata are up-to-date and pertinent, and that a prompt and knowledgeable support service is available. This dimension has three elements, namely:
    5.1      data accessibility,
    5.2      metadata accessibility, and
    5.3      assistance to users.

**Box 3: The Cascading Structure of the Data Quality Assessment Framework, DQAF July 2003, for the National Accounts Statistics: An Example**

*Using serviceability as the example of a dimension of quality, the box below shows how the framework identifies three elements that point toward quality. Within consistency, one of those elements, the framework next identifies three indicators. Specifically, for each indicator, focal issues are addressed through key points that may be considered in identifying quality.*

| Dimension | **4. Serviceability** |
|---|---|
| **Elements** | **4.1 Periodicity and Timeliness** |
| | **4.2 Consistency** |
| | **4.3 Revision policy and Practice** |
| **Indicators** | **4.2.1 Statistics are consistent within the dataset** |
| | **4.2.2 Statistics are consistent or reconcilable over a reasonable period of time** |
| | **4.2.3 Statistics are consistent or reconcilable with those obtained through other data sources and/or statistical frameworks** |
| **Focal Issues** | *i. The statistical series is internally consistent* |
| **Key Points** | • National accounts are internally consistent<br>– a set of consistent GDP estimates by activity and expenditure components is derived;<br>– if not, a statistical discrepancy between these estimates has not been large and has been stable over time;<br>– similar and consistent growth rates are obtained from the GDP series by activity and the GDP series by expenditure categories;<br>– total supply of goods and services matches the independently derived total uses;<br>– GDP estimates at current prices, volume measure, and (implicit) deflators are consistent within the "value = volume x price" framework. |

## V.  HOW IS THE DQAF USED?

The DQAF provides a structure for assessing existing practices against best practices, including internationally accepted methodologies. It has proved to be valuable for at least three groups of users.

➤ To guide IMF staff on the use of data in policy evaluation, preparing the Data module of [Reports on the Observance of Standards and Codes](#) (Data ROSCs), and designing technical assistance.

> **Within the IMF**, the framework is being used in a variety of circumstances. STA is using the DQAF in preparing Data ROSCs (see below), in conducting technical assistance, and in working with countries that wish to participate in the GDDS to prepare metadata, including their plans for improvement. With respect to Fund surveillance, the Data ROSC provides background information for the Article IV Consultation. In this connection, the DQAF provides a methodology that staff use to assess the quality of data provided to the Fund. In this sense, the DQAF is especially useful because it fosters an evenhanded approach to assessing quality across the very diverse range of countries that comprise the Fund's membership.

➤ To guide country efforts e.g., to prepare self-assessments.

> **National Statistical Office**. A Statistical office undertaking an internal assessment can use the framework. This assessment can be the basis for its own internal planning and for requesting and justifying additional resources.

➤ To guide data users in evaluating data for policy analysis, forecasts, and economic performance.

> **Financial market analysts and others**—researchers, for example—may find summaries included in a Data ROSC useful as a reference tool. To take one example, a financial market analyst might supplement the summary information provided in the Data ROSC with his/her own conclusions drawn from a specific dataset.

### Integrating the DQAF and Data ROSCs

From the inception of the ROSC initiative, the SDDS and the GDDS were used as the standards for the Data ROSC. For countries that have subscribed to the SDDS or are close to meeting the requirements for subscription, the SDDS serves as the standard against which the country's data dissemination practices are compared. In the case of others that have agreed to the preparation of a Data ROSC, the recommendations of the GDDS are used to guide this part of the assessment.

The early Data ROSCs focused on the disclosure elements of the standard—that is, the requirement to make information available to the public. However, experience showed that

the reports would be more useful if they also addressed the quality of the information provided. This need to focus more precisely on the quality of the data disseminated under the standard was addressed by integrating the methodology provided by the DQAF into the structure of the Data ROSC. It should, however, be emphasized that the DQAF is an assessment methodology and not a standard in itself.

The assessment methodology provided by the DQAF encompasses all of the dimensions covered in the SDDS and GDDS—including accessibility and integrity—and complements them in a number of respects. It is worth recalling that the SDDS and GDDS were established, respectively, to guide countries in the provision of data (and metadata) to the public (dissemination) and promote statistical capacity building. In the area of data quality, the SDDS and the GDDS call upon subscribers and participants, respectively, to provide a range of information to data users that could serve as monitorable proxies for the quality of the data disseminated. However, the DQAF takes a more structured approach by providing users with a methodology, based on directly observable features of quality, that "walks users through" what best practice would call for to assure quality in the collection, production, and dissemination of data. Application of the DQAF methodology helps identify those areas where further efforts are required of the country to reach an international "best practice"

While the IMF encourages countries to publish these assessments, publication—like participation—is voluntary. To date, over 100 Data ROSCs have been published for over 80 countries. These reports are available at: http://dsbb.imf.org/Applications/web/dqrs/dqrsroscs/

**Integrating the DQAF in the SDDS and GDDS**

Integrating the DQAF into the SDDS and the GDDS has two important implications for the effectiveness of statistical surveillance. First, it makes the SDDS and GDDS more effective in promoting the dissemination of quality information by organizing them under a data quality framework. Second, it also links SDDS and GDDS metadata, which describe the prevailing level of a country's statistical attainment, with the country's Data ROSC, which, in turn assesses the quality of the statistical processes underlying a country's data. These linkages enhance the usefulness of SDDS and GDDS metadata, along with the metadata from the Data ROSC, for the Fund's surveillance process, as these metadata could help strengthen discussions of the analytical implications of data deficiencies and potential remedies for these deficiencies.

**The application of the DQAF in technical assistance by the IMF**

The DQAF conveniently ties together, as a centralizing framework, the IMF's statistical work to strengthen member's practices through Data ROSCs, technical assistance (TA), and the data dissemination initiatives. Recent experience points to the key role of the DQAF in enhancing the prioritization and effectiveness of IMF TA in statistics. In this setting, some Data ROSC missions have provided diagnostics leading to TA, and others have validated statistical improvements resulting from earlier single topic or multisector TA missions. Also, Data ROSC and TA missions have facilitated countries' subscription to the SDDS or participation in the GDDS.

## VI.  CONCLUSION

The DQAF is a product of an intensive, iterative process of consultation with national statisticians, international organizations, and data users. The DQAF was endorsed by all the IMF's member countries, through the Executive Board of the IMF and it provides a good framework to improve transparency in data dissemination as well as a framework for TA.