

## Ethiopian Data Quality Assessment Framework (EDQAF)

Aberash Tariku Abaye;

Central Statistical Agency;

Addis Ababa, Ethiopia;

Email: [kaberash@yahoo.com](mailto:kaberash@yahoo.com)

Central Statistical Agency of Ethiopia (CSA) has developed the National Strategy for the Development of Statistics (NSDS). The NSDS is a road map for the statistical activity to be conducted in the five years. The NSDS have six themes and one of the themes refer to the development of EDQAF to improve the quality of statistical data by assessing the system and the output. The design of the EDQAF involved the participation of sector ministries. International experiences were also taken in to account in the EDQAF design.

EDQAF has four different phases. These phases are Phase 1: Initiation and Preliminary Investigation, Phase 2: Systems Assessment, Phase 3: Overall National Assessment, Phase 4: Reporting and Conclusion. EDQAF document defines the data quality dimensions in terms of output, process and Institutional Environment Quality Dimensions.

**Key words:** Data quality dimensions, data quality assessment phases.

## 1. Introduction

Good quality data is important for planning and to support decision-making. Monitoring and evaluation of Growth and Transformation Plan (GTP) and other millennium development goals requires good quality data. Assessing the quality of the data will ultimately result in improvements in the data quality.

The objectives of the Central Statistical Agency are to collect, process, analyze and disseminate the necessary socio economic and demographic statistical data through census, sample surveys, continuous registration and administrative recording system; and to provide technical guidance to government agencies and institutions in their endeavour to establish administrative recording, registration and reporting systems; and build the capacity required for providing directives and consultations in database creation and development of administrative records and registration systems. Looking at the powers and duties of CSA, one of the duties is to design and monitor the implementation of statistical recording and reporting systems to be followed by government agencies or institutions and other organizations. From this act and associated powers and duties, it can be deduced that CSA is mandated to develop and implement Ethiopian Data Quality Assessment Framework (EDQAF).

CSA designed National Strategy for Development of Statistics (NSDS) Covering period 2009/10 – 2114/15. Six strategic themes of which two are very relevant to quality issues. The six themes are; theme 1: implementation of the statistics law, theme 2 : developing data quality procedures, theme 3 : enhance advocacy and use of statistics, theme 4 : methodological improvements and statistical modernization, theme 5 : capacity development in the NSS, Theme 6: Relationship of NSDS to the monitoring and evaluation of Growth and Transformation Plan (GTP). Strategic Theme 1: Implementation of the Statistics Law describes about establishment of an NSS methodological and support unit in the CSA for quality assessment and NSS capacity building, Development of common standards, classifications & definitions for the NSS for consideration by the Council as legal decrees for official statistics, introduction of memoranda of understanding between the CSA and its NSS partners and Coordination of donor relations and statistical initiatives in the NSS. Strategic Theme 2: Develop data quality procedures states about developing a data quality assessment framework for Ethiopia, Development and support of ministry/agency statistical units in NSS partners, Strengthening of NSS quality and support unit in the CSA for quality assessment and NSS capacity building .To accomplish the duties given in the statistics act and implement the NSDS, an NSS methodological and support unit in CSA is established. The unit prepared common standards and classifications and also designed EDQAF.

## 2. EDQAF objectives

The EDQAF is designed to meet needs of data users by providing better quality data, providing data of known quality - quality reports and Identifying/labelling official statistics. The EDQAF is also designed to meet needs of data producers by Providing quality reports and Highlighting quality problems and proposals. It is also designed to meet needs of Other stakeholders/ funding agencies by Providing evidence of well organized statistical system and good quality data output.

The Target Outcome of EDQAF is more satisfied users, using better statistics to make more informed decisions. The Overall Objective is to introduce a comprehensive quality assessment program that summarises the quality of NSS data outputs for the benefit of both data producers and users, and that identifies quality problems and potential quality improvements and brings them to the attention of producers and senior managers for action. The Specific Objectives are to assess the quality of all NSS data outputs and of the systems that produce them, to identify quality problems, to propose quality improvements, to ensure that quality problems and potential improvements are brought to the attention of senior management, to review the extent to which quality problems have been addressed in the next assessment round, to provide producers and users with a quality summaries including quality scores by dimension and finally to provide quality scores enabling labelling of official statistics

### 3. The EDQAF development process

International standards and best practices are used in designing EDQAF. These standards are adapted to the Ethiopian situation. The Procedures/steps used are First draft EDQAF Developed by CSA, the draft EDQAF Distributed to NSS members for comments, stakeholder workshop conducted and now the final EDQAF is prepared and adopted by the statistics council. A user friendly soft ware for EDQAF data analysis is developed by CSA.

### 4. EDQAF basic assessment areas and scope

EDQAF is designed in such a way that it assesses both the system producing the statistical data and also the final out put. Two major areas of dimensions are defined; Output Quality Dimensions and Systems Quality dimensions. EDQAF Data Output Quality Dimensions are Relevance, Accuracy, Timeliness and Punctuality, accessibility and Interpretability, Coherence and Comparability. EDQAF Systems Quality (Process and Institutional) Dimensions are Methodological soundness, Human resource management, Standard operations, Data management and security, Quality assurance/control, Reporting burden, Mandate, Resource, performance and Quality Management, Integrity, Provider Transparency and Confidentiality.

A statistical out put data set produced by NSS is the target objects of EDQAF. NSS comprises CSA and other sector ministries producing statistical data. National level datasets produced by government agencies or institutions and that produce, or could produce government statistics are the scope of assessment .Regional, zonal, and woreda datasets which contribute to national datasets are also in the scope. Federal, Regional, zonal or woreda organisations are free to adopt and implement the EDQAF for self assessment.

### 5. EDQAF Assessment Phases

EDQAF assessment involves four assessment phases. These are Phase 1: Initiation and Preliminary Investigation, Phase 2: Systems Assessment, Phase 3: over all National out put Assessment, Phase 4: Reporting and Conclusion. Phase 1: Initiation and Preliminary Investigation involves setting up the assessment schedule and working relationships, undertaking preliminary discussions with data producer and reviewing documentation and establishing logistics of the assessment. Phase 2: Systems Assessment involves On site assessment at each of service delivery, woreda, zonal, regional and federal levels through which the data pass, It involves assessment of data collection, capture, processing and transmission procedures based on discussions with production staff, and review of metadata and other documentation at each level. Verification of (samples of) data received and transmitted at each level is also included. Phase 3: Overall National out put Assessment includes assessment of process, institutional and output quality based on detailed discussion with production staff at national office, review of metadata and other documentation at national level and review of results of Phase 2 assessment. Phase 4: Reporting and Conclusion involves preparation of assessment results, comprising quality summaries and descriptions of major quality problems and potential quality improvements , distribution and discussion of results with data producer, senior managers having oversight of data production process, and key data users, formally wrapping up assessment process and provisionally scheduling next assessment.

### 6. Questionnaire format of the EDQAF assessment

Standard questionnaire is designed for each phases of the assessment. Indicator related to the dimension defined in the questionnaire. Four levels of certification are designed for each indicator. Level 4 : quality statistics, Level 3: acceptable statistics, Level 2 : questionable statistics and Level 1: poor statistics. Level 4: Good indicates that the data/process/institution satisfies(s) all the quality requirements associated with the quality dimension. Level 3: Acceptable means the data/process/institution satisfies(s)

many of the quality requirements. Level 2: Questionable refers the data/process/institution satisfies(s) few of the quality requirements and Level 1: Poor explains the data/process/institution satisfies(s) none of the quality requirements, or cannot be assessed. Dataset will be labelled as a source for official statistics if and only if it has score of at least 3 for the accuracy and sound methodology, and a score of at least 3 averaged over all quality dimensions included in Overall National (Phase 3) Assessment.

### EDQAF Phase 2: Systems Assessment – Questionnaire format

| Element  | Indicator   | Benchmark Levels  |
|--|---|---|
| <i>Process Quality: Quality Dimension 7. Human Resource Management</i> |   |   |
| <b>7.1 Human Resource Management</b>                                   | a) Does the organizational structure clearly identify the positions that have responsibility for data production activities | 4. All data production responsibilities are clearly identified                |
|  |   | 3. Most data production responsibilities are clearly identified.              |
|  |   | 2. A few data production responsibilities are identified.                     |
|  |   | 1. No data production responsibilities are identified.                        |
|  | b) Are the staff positions dedicated to data production activities all filled?  | 4. All staff positions dedicated to data production are filled.               |
|  |   | 3. Most staff positions dedicated to data production are filled.              |
|  |   | 2. Only a few of the staff positions dedicated to data production are filled. |
|  |   | 1. There are no staff positions dedicated to data production per se.          |

#### 7. Conclusion

Good quality data are results of good quality statistical process operating with in a good quality institutional environment. A data quality assessment frame work which assesses the process, the environment and finally the out put data is expected to increase quality of data and thus the implementation of EDQAF is expected to result in improvement of the data produced by NSS.

## Reference

CSA (2009), Federal Democratic republic of Ethiopia, National Strategy for the Development of Statistics (NSDS)

CSA (2005), Federal Democratic republic of Ethiopia, Statistics Proclamation No 442/2005

CSA (2011), Federal Democratic republic of Ethiopia, Ethiopian Data Quality Assessment Framework (EDQAF)