FAO Statistics and Data Quality Assurance Framework

Endorsed by the Data Coordination Group (DCG)
Last revised (June 2023)
Contents

1. Introduction ........................................................................................................................................ 4

2. The FAO Statistics and Data Quality Assurance Framework ................................................................. 5

2.1 Scope and typology .............................................................................................................................. 7

2.2 FAO SDQAF definition of Quality ...................................................................................................... 8

➢ Relevance ........................................................................................................................................ 9

➢ Accuracy and Reliability ..................................................................................................................... 9

➢ Timeliness and Punctuality ................................................................................................................. 9

➢ Coherence and Comparability ........................................................................................................... 9

➢ Accessibility and Clarity .................................................................................................................... 9

2.3 FAO SDQAF principles ...................................................................................................................... 10

Institutional environment...................................................................................................................... 10

FAO SDQAF Principle 2: High and Continuous Commitment to Quality .................................................. 11

FAO SDQAF Principle 3: Adequacy of Resources .................................................................................. 11

FAO SDQAF Principle 4: Data Protection and Statistical Confidentiality ............................................... 12

FAO SDQAF Principle 5: Development and use of international standards .............................................. 12

FAO SDQAF Principle 6: Cooperation with data providers ..................................................................... 13

FAO SDQAF Principle 7: Coordination with other international organisations producing statistics ...... 13

Statistical Processes .............................................................................................................................. 14

FAO SDQAF Principle 8: Suitable and trustworthy data sources ............................................................ 14

FAO SDQAF Principle 9: Sound methodology and appropriate statistical procedures ............................. 14

FAO SDQAF Principle 10: Cost-effectiveness ......................................................................................... 15

FAO SDQAF Principle 11: Non excessive burden on respondents ......................................................... 16

Statistical Outputs .................................................................................................................................. 16

FAO SDQAF Principle 12: Relevance ..................................................................................................... 16

FAO SDQAF Principle 13: Accuracy and Reliability .............................................................................. 17

FAO SDQAF Principle 14: Timeliness and Punctuality .......................................................................... 17

FAO SDQAF Principle 15: Coherence and Comparability ..................................................................... 17

FAO SDQAF Principle 16: Accessibility and Clarity .............................................................................. 18

2.4 FAO SDQAF Compliance mechanisms ............................................................................................ 19

Procedures for evaluating existing statistical activities and data platforms .......................................... 19

Procedures for evaluating new data collections .................................................................................... 19

Procedures for evaluating the use of new non-statistical data sources in the production of statistical outputs ...................................................................................................................... 20

Leadership and coordination .............................................................................................................. 20

Implementation plan ............................................................................................................................ 20

Annex 1. Correspondence between FAO SDQAF principles and Principles Governing International Statistical Activities ................................................................. 21

Annex 2. Correspondence between FAO SDQAF principles and QAF principles of other international organisations .................................................................................................................. 26

Annex 3: Correspondence between the FAO SDQAF and the UNECE Big Data Quality Framework .......... 29

Annex 4. Glossary ................................................................................................................................... 35
1. Introduction

High quality data and statistics on food and agriculture provide the foundation for evidence-based policy making both for national governments and the international community. For FAO, sound data and statistics play a critical role in designing and targeting policies to reduce hunger, malnutrition and rural poverty, improve food systems productivity, sustainability and resilience, and combat climate change, biodiversity loss and inequality. They are also essential in monitoring progress towards national, regional and international development goals and targets. It is therefore imperative that the data and statistics produced by FAO are of the highest possible quality.

In 2014, FAO released its first Statistics Quality Assurance Framework (SQAF) which has been used to guide best practices in managing statistics quality at the level of FAO statistical outputs, statistical processes and institutional environment. Several of the SQAF recommended best practices are now part of FAO culture and have been consolidated in FAO statistical standards, policies and governance mechanisms.

In eight years, the environment of statistics has, however, evolved. On one hand, data demands for timely food and agricultural statistics have never been so great. The data needs of the 2030 Agenda for sustainable development are four times those of the Millennium development goals. And, recent global events such as the Covid-19 pandemic, the desert locust and the war in Ukraine have also increased pressures for the provision of timely actionable insights to mitigate the impacts of such events on food security and agri-food systems. On the other hand, in the last decade, the relationship between data and statistics has become even closer with the proliferation of new or alternative data sources such as big data, earth observation data and administrative data, which can be used in the production of statistics alongside traditional censuses and surveys. The fact that these alternative data sources, the technologies to use them and their related data science skills are becoming more accessible offers opportunities like never before for the production of more timely statistics, but comes hand-in-hand with concomitant risks that must be managed.

To address this changing environment, FAO has introduced since 2020 some changes to better integrate and improve FAO data and statistics. Improvements were introduced in FAO data and statistics governance mechanisms, namely the nomination of an Executive Data Champion, the creation of a Data Coordination Group and the expansion of the mandate of FAO Inter-Departmental Working Group on Statistics (IDWG-Statistics) to data and statistics. A Data lab for statistical innovation was also created to develop FAO’s capabilities in data science and further explore the use of non-traditional data sources in statistical production. FAO Policies on Data Protection and Intellectual Property Rights were also promulgated in 2022-23, which impacted all sensitive, confidential and strictly confidential data processed by FAO, including those used in the production of statistics or data for statistics. The revision of the SQAF towards a harmonized quality assurance framework for data and statistics was therefore a necessary step to ensure that the use of alternative data sources to fill existing data gaps is properly managed and the resulting statistical products satisfy the highest quality standards.

The resulting new FAO Statistics and Data Quality Assurance Framework (SDQAF) will complement the efforts being made to increase coordination and to improve consistency of the overall statistical
programme. It will also help to ensure that sound and internationally accepted statistical concepts and definitions are adhered to, standard methodologies are adopted, and that accurate, high quality, timely and accessible data are produced by FAO.

This corporate quality assurance framework for data and statistics is particularly important within FAO decentralized statistical system, where each FAO technical department carries out its own statistical programme of work and maintains ownership of its data. In such system, and without the proper coordination mechanisms and overall framework, common standards and methodologies are applied on a good-will basis and not on a consistent corporate scale.

It is worth noting that, similarly to other international organizations, FAO statistics are mainly based on data supplied by national statistical agencies, other international organizations, or, to some extent, other trustworthy sources. Consequently, their quality depends on the quality of inputs received. This document provides an overall framework to ensure the quality of FAO statistical outputs, and the soundness of processes and governance mechanisms put in place for their production. The quality elements specific to different typologies of input data used by FAO are considered in dedicated and complementary corporate statistical standards.

2. The FAO Statistics and Data Quality Assurance Framework

The main objective of the FAO Statistics and Data Quality Assurance Framework (FAO SDQAF) is to guide the Organization in ensuring the availability of high-quality international statistics on food and agriculture for decision-making. As shown in Figure 1, the framework consists in a series of principles to adhere to at the institutional, statistical process and statistical output level, in order to ensure the quality of FAO statistics. The quality of statistics itself is defined in the framework, through various quality dimensions, in order to guarantee common understanding of what it is to achieve.
FAO Statistics and Data Quality Assurance Framework

 Ensuring the availability of high quality international statistics on food and agriculture for decision-making

In addition, each principle is accompanied by corresponding key implementation modalities, which provide information on how FAO manages and achieves each principle. Finally, the framework describes the main mechanisms in place to assure and monitor FAO’s compliance to its SDQAF.

The FAO SDQAF principles encompass the **Fundamental Principles of Official Statistics of the United Nations Statistical Commission**\(^1\), as well as the **Principles Governing International Statistical Activities** endorsed by the Committee for the Coordination of Statistical Activities (CCSA)\(^2\). Correspondence between the FAO SDQAF and the Principles Governing International Statistical Activities is reported in Annex 1. In developing to its new SDQAF, FAO also examined the quality assurance frameworks of other key international organizations (European Central Bank, Eurostat, IMF, OECD). Correspondence between their respective principles were documented in Annex 2. Finally, the FAO SDQAF also takes into account the **Suggested Framework for the Quality of Big data**\(^3\) developed by the UNECE Big data quality Task Team (see correspondence in Annex 3) and the CCSA’s **Recommended practices on the Use of Non-Official Sources in International Statistics**\(^4\).

The implementation of the SDQAF is of key importance to strengthen and sustain FAO’s reputation and credibility as a centre of excellence in agricultural data and statistics. Successful application of the principles and their key implementation modalities described in this document will contribute directly to improving the confidence of users in FAO statistical outputs.

---

2.1 Scope and typology

The framework is intended to apply to all data and statistical activities, including statistical development activities implemented by FAO and/or conducted by any FAO employees. This comprises data and statistics collected, produced and released by FAO as part of its statistical programme, either disseminated as global public goods or used by FAO to inform the implementation of its programmes.

It does not include human resource, or financial and ICT infrastructure management data (operational data related to running the organization) as they are not specific to FAO’s data production functions laid out in the article I of its Constitution. For the purposes of the framework, a statistical activity is one in which the primary focus is acquiring, processing, storing, analyzing, and disseminating statistical data. It does not include the use of statistics. As for statistical development activities, they generally refer to the establishment and/or improvement of statistical methods, standards and procedures used in the production and dissemination of statistics.

As described in the Section 2.2., quality in the statistics context is ultimately measured in terms of the quality of the final statistical outputs, even though this can only be assured through a proper management of quality at the institutional and statistical process level. At the institutional level, this consists in acting upon and controlling the institutional and organizational factors that have a significant influence on the Organization’s effectiveness and credibility to assure high-quality statistical processes and outputs.

At the statistical process level, the quality is managed throughout the production cycle that leads to the statistical outputs by ensuring that international standards, guidelines and best practices are applied, in a cost-effective matter, at the various stages of the planning, development, production and dissemination of the statistical outputs. In general, FAO statistical processes follow the Generic statistical business process model (GSBPM) v5.1 developed by UNECE High-Level Group for the Modernisation of Official Statistics (HLG-MOS).

However, the typologies of FAO input data and statistical outputs influence greatly how the GSBPM statistical processes and sub-processes are applied. As shown in Figure 2, most of FAO data sources are secondary data sources that can be qualified as statistical or non-statistical. Statistical data sources include data that were produced for a statistical purpose (e.g. collected through a representative surveys) or produced and validated as statistics from their data providers (e.g. official national statistical estimates produced by a country). Non-statistical data sources (also referred to as non-traditional data sources) comprises data that were not primarily collected for a statistical purpose but could be used in compiling official statistics (e.g. administrative data, big data such as remote sensing information).

---

5 https://www.fao.org/3/mp046e/mp046e.pdf
6 https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1
Figure 2: Typologies of FAO input data and statistical outputs

<table>
<thead>
<tr>
<th>FAO INPUT DATA</th>
<th>FAO STATISTICAL OUTPUTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical data sources</strong></td>
<td><strong>National estimates</strong> (including disaggregated estimates), which may have been modified by the FAO statistical process</td>
</tr>
<tr>
<td>- Official data from national recognized agencies or provided by officially nominated focal points(^2), collected through FAO questionnaire or obtained directly from official national statistical products, regardless of the data sources used at national level.</td>
<td><strong>New National estimates calculated by FAO</strong> (derived variables; complex indicators calculated using different input variables; disaggregated estimates)</td>
</tr>
<tr>
<td>- Data from international organizations (official or estimated), obtained through data sharing agreement or directly from open-source statistical products, regardless of the data sources used by the international organizations.</td>
<td><strong>Aggregates of the collected input data</strong> (global and regional aggregates)</td>
</tr>
<tr>
<td>- Primary statistical data (e.g. microdata), collected by FAO through probabilistic and representative sample surveys or obtained through microdata sharing protocols with recognized national or international data producers</td>
<td><strong>Aggregates referred to the new variables/indicators</strong> produced by FAO (global and regional aggregates)</td>
</tr>
<tr>
<td>- Administrative data, collected by FAO in the context of its operations or obtained through data sharing agreements with the data owner</td>
<td><strong>Aggregates obtained directly by regional or international institutions</strong></td>
</tr>
<tr>
<td>- Big data (raw data or pre-processed), obtained either through a cooperation with the data owner or accessed freely through web scraping (e.g. earth observation data, mobile phone data, data web-scraped from social media...)</td>
<td><strong>Microdata files</strong> (e.g. FIES) (after data treatment and confidentiality preservation procedures by FAO or agencies providing data)</td>
</tr>
</tbody>
</table>

\* More detailed definitions have been included in Annex 4.

These typologies for example simplifies the activities related with the design and creation of frames and samples as well as data processing when national official estimates provided by countries or international organizations are of good quality and re-disseminated with minimal data manipulation and validation. On the other hand, the use of secondary data increases the complexity of the data integration process, especially for non-statistical data sources, and makes the validation of outputs more difficult as FAO doesn’t necessary have a direct access to the primary data or respondents.

### 2.2 FAO SDQAF definition of Quality

Quality is a multi-faceted and subjective concept. The International Organisation for Standardization (ISO) defines quality as “the totality of features and characteristics of a product or service that bears on its ability to satisfy stated or implied needs” (ISO No 8402; 1986, 3.1).

The most important quality characteristics depend on user perspectives, needs and priorities, which vary across groups of users. For this reason, the major challenge is to achieve a compromise among the needs of the various possible users (current and potential) in order to produce and disseminate

\(^2\) Often referred to as secondary data
statistical outputs that satisfy the most important needs given constraints concerning available resources.

Taking stock of the work already done by several international organizations (Eurostat, ECB, IMF and OECD) in this area, the definition of quality in the context of statistical outputs, which has been tailored to the FAO framework, encompasses five quality dimensions, as described below.

FAO defines quality in statistics as *the degree to which its statistical outputs fulfil requirements* and the following quality dimensions are taken into account:

- **Relevance**
  
  - Relevance is the degree to which statistics meet the current and potential user needs.

- **Accuracy and Reliability**
  
  - Accuracy refers to the closeness of estimates, to the true values that statistics were intended to measure.
  
  - Reliability refers to the closeness of the initial estimates to the subsequent or final estimates.

- **Timeliness and Punctuality**
  
  - Timeliness is the speed of dissemination of statistical outputs - i.e. the lapse of time between the end of a reference period (or a reference date) and the dissemination of the statistical outputs.
  
  - Punctuality refers to the possible time lag existing between the actual delivery date of statistical outputs and the target date when they should have been delivered, for instance, with reference to dates announced in an official release calendar or previously agreed among partners.

- **Coherence and Comparability**
  
  - Coherence is the adequacy of the statistical outputs to be meaningfully combined in different ways and for various uses.
  
  - Comparability refers to the extent to which differences between different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistical characteristics.

- **Accessibility and Clarity**
  
  - Accessibility is defined as the ease, the set of conditions and the modalities by which users can obtain data.
  
  - Clarity refers to the availability of adequate documentation: whether data are accompanied with appropriate metadata, illustrations such as graphs and maps, whether information on their quality are also available (including limitation in use), and the extent to which additional assistance is provided.

---

8 Annex 2 reports the mapping of FAO SDQAF principles with the principles of other organizations' Quality Assurance Frameworks
2.3 FAO SDQAF principles

The FAO SDQAF principles and their key implementation modalities are set out in three areas: Institutional Environment, Statistical Processes and Statistical Outputs.

The first priority of the FAO statistical system is to meet user needs providing them high quality statistical outputs. Given that such level of quality is intrinsically dependent on the soundness of governance mechanisms and processes in place for the statistical production, the principles related to the Institutional Environment are described first in the FAO SDQAF. The second part of the FAO SDQAF is devoted to principles related to Statistical Processes, given that high quality statistical outputs can only be produced by high quality statistical processes. The third part of FAO SDQAF is dedicated to the principles which relate to Statistical Outputs.

Institutional environment

Institutional and organizational factors have a significant influence on the Organization’s effectiveness and credibility to assure high-quality statistical processes and outputs. Compliance with the following principles ensures that FAO statistics are produced in an environment that promotes and supports high-quality standards in the implementation of its processes and outputs.

FAO SDQAF Principle 1: Professional independence and impartiality
FAO statistics are produced and disseminated on the basis of scientific and professional independence from other policy, regulatory or administrative departments and bodies, as well as from private sector operators. This is done in an objective, impartial, professional and transparent manner.

Key implementation modalities
1.1. The selection of data sources, as well as the collection, compilation and dissemination of FAO statistics are undertaken according to strictly scientific and professional statistical considerations, and on an objective and impartial basis.
1.2. Information on the sources, methods and procedures used in statistical processes is publicly available.
1.3. FAO publishes in advance its calendar of release and explanations of any delays are publicly available.
1.4. Statistical releases are issued separately from political or policy statements.
1.5. Press releases on statistical outputs are objective, impartial and accessible to a wide range of audiences.
1.6. Criticisms, erroneous interpretation and misuse of statistics are immediately and appropriately addressed. Errors discovered in published statistics are corrected as quickly as possible and actions are taken to avoid the occurrence of similar errors.
1.7. Advanced notice is given on major revisions or changes in methodologies.
1.8. Decisions about FAO statistical work programmes are made publicly available through various channels, including through official reports presented to FAO governing bodies, FAO regional commissions on agriculture statistics and the UN Statistical Commission.
1.9. All documents for and reports of statistical meetings, statistical capacity building initiatives, and technical assistance projects are publicly available on FAO website.
**FAO SDQAF Principle 2: High and Continuous Commitment to Quality**

FAO is highly committed to quality. The strengths and weaknesses of FAO statistics are systematically and regularly monitored to continuously improve the quality culture within the organization and the quality of both processes and products.

Key implementation modalities

1. FAO data quality framework for statistical outputs and statistical standards are defined and publicly available. They are regularly updated to take into account the changing environment and new best practices.

2. All personnel are individually responsible for the quality control and assurance of the data they process. The FAO Data Coordination Group, the FAO Inter-Departmental Working Group on Statistics, the FAO Chief Statistician, and the FAO SDQAF with its methods, tools and implementation plan play a key role in quality management and leadership within the Organization.

3. Procedures are in place to regularly monitor the quality of statistical production processes, including the integration of data from multiple data sources, and identify best actions to address data quality challenges.

4. Product quality is regularly monitored, assessed with regard to possible trade-offs, and reported according to the FAO SDQAF.

5. There are regular and thorough review of FAO key statistical outputs, involving external experts and users where appropriate, such as users’ consultation, peer reviews or external evaluations.

6. The FAO training programme includes training courses designed to support the implementation of its data quality framework.

7. FAO engages in capacity development activities targeting data providers, in particular national statistical agencies, to improve the quality of their data and statistics.

8. FAO has a corporate risk management tool that helps monitor, mitigate and manage the risks associated to the quality of its data and statistics.

**FAO SDQAF Principle 3: Adequacy of Resources**

The resources available to FAO statistical programmes are sufficient to meet data and statistics requirements. The establishment of a corporate resources allocation and monitoring mechanism is an essential way of ensuring that human, financial and technical resources are adequate.

Key implementation modalities

1. FAO has a dedicated budget for FAO statistics, both in terms of staff and non-staff resources, to implement its statistical processes that is commensurate to priority food and agriculture statistics needs.

2. When appropriate, FAO takes actions to mobilize additional resources, from corporate or extra-budgetary sources, for statistical infrastructure enhancement projects, research and development, new data demands and capacity development programmes supporting national food and agriculture statistics systems.

3. Governance and procedures are in place to monitor, document and review the way in which resources dedicated to statistics are used and prioritized as well as to encourage collaboration and joint activities where appropriate.
**FAO SDQAF Principle 4: Data Protection and Statistical Confidentiality**

Any data used or processed by FAO, or transferred to a third party by FAO, is processed correctly and consistently throughout the data lifecycle, that is, from initial collection, storage to deletion. All data of confidential or sensitive nature (e.g. concerning people and legal entities, small aggregates or sensitive location data) are kept strictly confidential, used exclusively for statistical purposes, and processed in line with FAO Data Protection Policy.

**Key implementation modalities**

4.1. FAO has confidentiality and protection policies on data and statistics that are made available to the public, and regularly monitor the compliance of its data production processes with these policies.

4.2. Guidelines and instructions are provided to staff in order to assess the confidentiality level and associated risks of their data, and adequately protect confidential and sensitive data used in statistical production and dissemination processes.

4.3. Physical, technological, methodological and organizational provisions are in place to protect the security and integrity of FAO statistical databases and underlying raw data whereas appropriate.

4.4. FAO promotes the definition and adoption of formal agreements with data providers to support data exchange, which ensure that confidentiality is respected whilst also permitting proper data use and statistical analysis.

4.5. A framework describing methods and procedures for exchanging micro-data with data providers is adopted, in line with data protection policies of respective parties.

4.6. FAO applies the most advanced international standards and best practices in the treatment and dissemination of microdata or georeferenced statistics (which involve data anonymization and generalisation, use of statistical disclosure procedures and/or strict terms of use for microdata).

4.7. FAO requests the informed consent of respondents participating in surveys directly carried out by the Organization.

4.8. All FAO employees are bounded by a duty of confidentiality, including for confidential data that do not cease upon separation from service and are also required to comply with FAO Data Protection Policy.

4.9. In case of a data breach, FAO has a process in place to take timely mitigation action.

4.10. FAO supports national data providers in complying with their legal commitments concerning data confidentiality.

**FAO SDQAF Principle 5: Development and use of international standards**

FAO contributes to, and promotes, the definition and implementation of international statistical standards in all fields of work within FAO’s mandate. Standards for international statistics are developed on the basis of sound professional criteria, while also addressing issues of practical utility and feasibility.

**Key implementation modalities**

5.1. FAO systematically involves national statistical offices and other national, regional and international organizations for official statistics in the development and promulgation of methods, standards and good practices.

5.2. FAO systematically works with other organizations towards agreements on common statistical concepts, classifications, standards and methods.

5.3. FAO ensures that decisions on international standards on statistics are free from conflicts of interest and are perceived to be so.

5.4. FAO advises countries on implementation issues concerning international standards and monitors the implementation of agreed standards.
**FAO SDQAF Principle 6: Cooperation with data providers**

*In order to produce high quality statistics, FAO promotes cooperation with its various data providers, including undertaking capacity development initiatives aimed at improving the quality of statistics produced by Member Countries.*

**Key implementation modalities**

6.1. Formal data acquisition protocols and/or data sharing agreements are in place with Member Countries and other data providers for the provision of data and metadata to FAO. These protocols and agreements specify the role that the data plays in the final statistical outputs, explain the data access and exchange procedures between the parties, and are compliant with FAO Policies on Data Protection and Intellectual Property Rights, where applicable.

6.2. FAO encourages national statistical agencies to strengthen the coordination of their national statistical system, in particular to help identify the responsible authorities in charge of providing official data to FAO, facilitate data reporting and better coordinate statistical capacity development activities.

6.3. Effective collaboration mechanisms with data providers are established to discuss how the quality of input data and metadata as well as of the arrangements under which data are shared with FAO could be improved.

6.4. FAO advocates the application of Fundamental Principles of Official Statistics when FAO input data are collected by national statistical agencies.

6.5. Relevant national official source organizations are consulted as appropriate when non-official data are published to fill gaps and missing values in official time series.

6.6. FAO promotes bilateral and multilateral cooperation in sharing knowledge (e.g. supporting training, project and guideline development) with countries and regions to further develop national and regional statistical systems.

**FAO SDQAF Principle 7: Coordination with other international organisations producing statistics**

*FAO promotes coordination with other statistical organizations in the development of international statistical programmes.*

**Key implementation modalities**

7.1. FAO participates in international statistical meetings, and bilateral and multilateral consultations for the development of international statistical programmes, whenever necessary.

7.2. There’s a procedure in place to coordinate FAO’s participation in international statistical meetings and its contribution to international working groups.

7.3. FAO promotes coordination with other statistical organizations in the development of coordinated international statistical programmes, in order to avoid overlaps and duplication of work.

7.4. FAO promotes bilateral and multilateral cooperation in sharing knowledge with international organizations to further develop national, regional and international statistical system.
Statistical Processes

The way statistical processes are conducted directly affects the quality of statistical outputs. FAO is committed to continuously improving the quality of its statistical processes, while ensuring their cost-effectiveness and that response burden on data providers is an acceptable level.

**FAO SDQAF Principle 8: Suitable and trustworthy data sources**

Data for statistical purposes may be drawn from all types of sources, be they official statistics provided by Members, international statistics produced by other organizations, data collected from statistical surveys or non-statistical data, such as big data.

*FAO chooses its data sources according to strictly professional considerations. The availability of official data as well as the accuracy, the relevance, the independence, the stability and the expected availability over time of the input data are among the most important factors considered by FAO when selecting its data sources.*

**Key implementation modalities**

8.1 FAO data sources are chosen using strictly professional considerations to ensure high quality statistical outputs.

8.2 Official statistics is typically the best source of information for the compilation of FAO statistics, however the Organization can refrain from publishing official data in cases of reasonable doubts on their quality.

8.3 The use of non-official sources is considered only when national official sources are not available or are of proven poor quality.

8.4 FAO uses non-official data sources in the generation of statistics only after a thorough quality assessment of the available official national and international statistics and of the alternative input data. When evaluating alternative data sources, priority is given to data produced by national statistical agencies or officially provided to the other international organizations by these agencies.

8.5 A procedure is in place to assess the adequacy of a non-official/non-statistical data source, including of its data, underlying data-generation process, metadata, and whereas appropriate of the institutional and business environment of its data provider. This assessment is done with respect to the intended uses of the data source in FAO statistical production processes and its impact on the quality of the resulting statistical outputs.

8.6 Methodology and criteria for selecting and assessing the adequacy of non-official/non-statistical data sources are made publicly available and equally accessible to all users.

8.7 When using non-official/non-statistical sources to fill gaps or to improve data quality and comparability of FAO statistics, FAO strictly follows internationally agreed recommended practices when they exist.

**FAO SDQAF Principle 9: Sound methodology and appropriate statistical procedures**

In developing and compiling data and statistics, FAO uses sound methodologies, and applies effective and efficient procedures throughout the data production cycle. FAO defines its methodological framework in accordance with established and internationally agreed scientific principles, and ensures its implementation across all its data production processes. FAO relies on highly competent staff to develop and successfully apply such sound methodology.

**Key implementation modalities**

9.1 The FAO methodological framework complies with international standards, guidelines and good practices, while constantly striving for innovation.
9.2 Procedures are in place to ensure that standard concepts, definitions, classifications and sound methodologies are consistently applied across FAO and its statistical processes.
9.3 Statistical methods are implemented in efficient, robust and standardized IT systems.
9.4 FAO has endorsed the ISI’s Declaration on Professional Ethics and encourages all its staff involved in statistical production to adopt high-standard professional values and ethics in statistics.
9.5 The FAO statistical system recruits highly competent staff from relevant disciplines and with appropriate qualifications.
9.6 FAO staff participate in training courses, on-the-job capacity development activities, conferences, and international meetings, and are encouraged to publish their work in peer-reviewed professional journals, to ensure that their methodological knowledge is up-to-date and shared globally, within the limitations of budget constraints.
9.7 FAO cooperates with the scientific community to improve methods and standards related to official statistics in general and food and agricultural statistics in particular.
9.8 Procedures are in place to decide whether a new statistical process should be implemented and how this process should be designed and applied.
9.9 The statistical procedures applied in data collection, processing, analysis, validation and dissemination are regularly documented, monitored, revised and disseminated.
9.10 Metadata related to statistical processes are managed throughout the statistical processes and disseminated, as appropriate.
9.11 Data collection forms and questionnaires used by FAO in its data collection processes follow corporate standards and are systematically tested prior to data collection.
9.12 Data revisions follow standard, well-established and transparent procedures.
9.13 When new non-traditional data sources are used for statistical purposes, their treatment and processing (deriving new variables, creating new data files, integrating data sources, etc.) follow appropriate specific procedures for controlling and assuring the quality of the output data based on internationally agreed recommended practices where they exist.
9.14 FAO Data transformation and analysis are based on theoretical principles and not dependent on the system that is performing them. For example, the residuals of a regression should be the same regardless of the analytical system performing the regression.

**FAO SDQAF Principle 10: Cost-effectiveness**
Efforts are made to continuously identify new and improved cost-effective approaches for carrying out FAO statistical processes. The establishment of a corporate monitoring mechanism is an essential way of ensuring that the most cost-effective approaches are adopted at FAO.

**Key implementation modalities**
10.1 FAO promotes and implements standardized solutions that increase cost efficiency and effectiveness.
10.2 Information and communication technology (ICT) tools are used to optimize the different phases of the statistical processes, including the storage and processing of large non-traditional input data sources (e.g. big data and geospatial information).
10.3 Cost-benefit and risk analysis including maintenance burden is carried out before new, or substantially enhanced, statistics processes are developed and integrated into regular production.
10.4 Governance and procedures are in place to monitor, document and review the way in which resources are used and to encourage collaboration and joint activities where appropriate.

---

**FAO SDQAF Principle 11: Non excessive burden on respondents**

FAO adopts strategies for monitoring, limiting and reducing respondent burden over time. The Organization appropriately chooses sources including non-traditional data sources (e.g. big data) and data collection methods, to minimise the reporting burden for data providers. The challenge of meeting user needs is balanced with managing the burden placed on respondents.

**Key implementation modalities**

11.1 FAO requests to data providers are limited to only what is absolutely necessary.
11.2 Mechanisms are in place to assess the necessity of undertaking new data collection.
11.3 Existing statistics are regularly monitored to ascertain whether the respondent burden can be further reduced.
11.4 Data sharing and overall coordination within FAO is generalised to avoid duplication of processes and to identify where the burden on data providers can be reduced.
11.5 FAO is committed to make the provision of data or data sharing to FAO easier for countries or other constituencies.
11.6 FAO coordinates with other organizations for data sharing, applies open data principles and has clear licensing policy to facilitate the use of its data by other organizations.
11.7 Where appropriate, FAO collects data directly from other international organizations, in order to reduce the response burden on national organizations.
11.8 Proactive efforts are made to improve the statistical potential of new emerging data sources and to limit response burden of national data providers (countries).

**Statistical Outputs**

With regard to its statistical outputs, FAO recognizes Principle 1 of the Principles Governing International Statistical Activities, endorsed by the Committee for the Coordination of Statistical Activities (CCSA), which states that “High quality international statistics, accessible for all, are a fundamental element of global information systems”. More specifically, the FAO SDQAF includes the principles below.

**FAO SDQAF Principle 12: Relevance**

FAO data and statistics meet user needs. In order to identify user needs, and to assess if they are met, FAO has regular and ongoing dialogues with its data users, and the information collected by such dialogues are reflected in data and statistical activities.

**Key implementation modalities**

12.1 Procedures are in place to identify and consult the key internal and external users of FAO data and statistics, in order to identify new or emerging requirements, and to monitor the relevance and usefulness of existing statistics.
12.2 Key users are consulted by FAO when the development of new data products, or the application of substantial changes to existing data are planned.
12.3 The FAO Statistical Programme of Work is periodically revised, in order to take into account changing user priorities.
12.4 User satisfaction is monitored regularly and follow up actions take place.
**FAO SDQAF Principle 13: Accuracy and Reliability**

FAO statistics accurately and reliably portray reality. FAO aims to minimize the variability as well as to the bias associated to its statistical outputs and estimates. In international statistics, lack of accuracy is generally due to the validity of data sources and errors which occur during the final production process. The quality of data sources is a key feature assessed by FAO when using non-statistical data sources as input for the production of statistical outputs. When revisions are made, FAO has a policy specifying the number and planned timing of revisions, and systematic analysis of revisions are undertaken to assess reliability.

Key implementation modalities

13.1 Source data, integrated data, intermediate results and statistical outputs are regularly assessed and validated.

13.2 Errors are measured and systematically documented. Statistical outputs based on the use of non-traditional data sources, such as big data, should come along with an assessment of selectivity and all other potential errors that can affect their accuracy.

13.3 Methods and tools for preventing and reducing errors are in place and used.

13.4 The revision policy adopted by FAO is made publicly available.

13.5 Revisions are regularly analysed in order to improve statistical processes.

---

**FAO SDQAF Principle 14: Timeliness and Punctuality**

FAO statistics are timely and punctual. FAO strives to produce as timely data and statistics as possible as non-timely statistical outputs can easily lose relevance. In doing so, the trade-off between timeliness and accuracy is taken into account. FAO recognizes that punctuality provides added value for users.

Key implementation modalities

14.1 FAO strives to produce as timely statistics as possible, without unduly compromising other quality features, and systematically works on the improvement of the timeliness of its statistics.

14.2 The periodicity of statistics takes into account user requirements as much as possible.

14.3 FAO agrees on a reporting timetable for the submission of data with data providers, which enables FAO to define a release calendar. FAO then monitors the compliance of data providers with the reporting timetable.

14.4 Preliminary results of acceptable aggregate accuracy and reliability can be released when considered useful.

14.5 The FAO statistics release calendar is publicly available, and FAO data releases are planned accordingly.

14.6 Delays in dissemination time schedule are announced in advance, explained and communicated together with a new release date.

---

**FAO SDQAF Principle 15: Coherence and Comparability**

FAO statistics are consistent internally, and over time. FAO statistics are internationally comparable and can be used in conjunction with statistics provided by other statistical organizations. As comparability is a key issue for international statistics, the adoption of standards on concepts and definitions is fundamental to achieving comparability. FAO fosters the development and implementation of international standards on food and agriculture statistics.

Key implementation modalities

15.1 FAO promotes cooperation and the exchange of knowledge between individual statistical programmes and domains to assure that its statistics are internally consistent.

15.2 FAO statistics are kept comparable over a reasonable period of time and between geographical areas. Effects of changes of the input data sources or of methodologies on final estimates are assessed and appropriate information is provided to users.
15.3 FAO Statistics are based on internationally agreed standards on concepts, definitions and methods, and enable international comparisons to be drawn. Conceptual differences, if any, are explained clearly to public. Significant changes in the phenomena to be measured are reflected by appropriate changes to concepts, classifications, definitions and target populations.

15.4 Statistics from different data sources, and with different time periodicities, are compared and any differences are explained (and reconciled).

**FAO SDQAF Principle 16: Accessibility and Clarity**

FAO statistics are easily accessible to all users on an impartial basis, presented in a clear and understandable format, and accompanied by relevant supporting metadata. Ensuring that users are able to easily access statistical outputs is a fundamental part of FAO statistical processes. The Organization works to continuously improve the accessibility of its data and statistics, taking into account users’ needs, new possibilities offered by IT developments and open data principles. Metadata accompanying data dissemination improve the clarity and interpretability of statistics.

**Key implementation modalities**

16.1 All users have equal and simultaneous access to statistical releases. Key statistical outputs are freely available for users, accessible in open data format and accompanied by clear terms of use.

16.2 On the basis of users’ needs and preference, FAO statistics are disseminated in various forms (e.g. databases, tables, publications, data visualizations...) and through appropriate channels (e.g. websites, social media, on-line fora...) which can be accessed through a unique entry point on FAO website.

16.3 FAO promotes dissemination-related applications and tools allowing users to generate and customize their own outputs (data, tables, graphical representations) according to their needs.

16.4 Access to microdata collected by FAO is allowed for research purpose and is subject to specific rules and protocols.

16.5 Users’ experience and needs with regards to the accessibility of FAO statistical outputs are regularly assessed and considered to improve FAO’s dissemination practices.

16.6 FAO statistics are disseminated together with relevant metadata presented in a standard format that facilitates proper interpretation, meaningful comparisons and machine-processing in line with the open data format.

16.7 Concepts, definitions and classifications, as well as data sources, data acquisition, data collection and processing procedures used, and the quality assessments carried out, are properly documented, and this information is publicly accessible.

16.8 When using non-official data sources in the production of the statistical outputs, their associated metadata clearly document the divergence in concept, definitions and data coverage of input data, the methods used to treat, process and obtain the final statistical outputs as well as the limitations to the uses of the statistical outputs caused by the complexity of the data sources.

16.9 In the dissemination of international statistics, FAO gives credit to the original source of data, and uses agreed quotation standards when re-using statistics originally collected by others.
2.4 FAO SDQAF Compliance mechanisms

In order to ensure the compliance of statistical processes, and corresponding outputs with its product and process quality principles, FAO has developed a series of tools and procedures to evaluate both existing and new statistical processes as well as the use of new data sources in the production of FAO statistical outputs. In addition, a set of institutional activities are being established and promoted so that FAO meets all the principles, which relate to the institutional environment.

**Procedures for evaluating existing statistical activities and data platforms**

The procedure for evaluating existing statistical processes is two-fold and consists of a combination of quality-related self-assessment and audit. Self-assessment is applied to most on-going statistical activities and their outputs, while usage of audit procedures is limited to complex and key corporate statistical processes and data platforms.

For the **self-assessment**, the officer responsible for the statistical process evaluates the adherence of the process and its outputs to the principles stated in the SDQAF, the key implementation modalities referred in the document, and the best practices promoted by FAO's Corporate Statistical Standards. The self-evaluation is completed through the FAO Quality Assessment and Planning Survey (QAPS), which is an internal survey implemented by the FAO Office of the Chief Statistician on a biannual basis. The QAPS relies on a questionnaire allowing to report on quality-related aspects of statistical activities at all phases of the statistical production process and by all principles included in the SDQAF. Self-assessment results are summarized in the QAPS final report, highlighting strengths and weaknesses of statistical processes and related outputs, and providing the basis to implement tailored improvement actions.

**Audit** is carried out by a team of auditors (2 or 3 internal or external experts as needed) who study the supporting documentation on the statistical activity, and then meet the officer responsible of the process to further investigate the identified issues. A checklist facilitates the completion of the task, the results of which are summarized in a final report providing recommendations for improvement actions to be implemented by the responsible officer.

**Procedures for evaluating new data collections**

When new statistical activities involving new data collection are planned, the assessment process aims to evaluate whether conditions are met in setting up the new process in terms of relevance (user needs) and respondent burden and overlap with other existing statistical activities. As a result, information concerning relevance, data sources and response burden is collected. Basic information on how the new process would be carried out, and a cost-benefit analysis, are also produced. This framework provides the foundation of the approval process. When the new statistical activity will involve surveys or other interactions with data providers to obtain personal data, a Privacy Self-Assessment and a Data Protection Impact Assessment should be completed.
**Procedures for evaluating the use of new non-statistical data sources in the production of statistical outputs**

The use of new non-traditional data sources in the production of statistical outputs is also carefully examined by the Organization. The assessment process can include a due diligence process on the data providers (if the data source is not freely available) and a thorough analysis of the relevance, validity, comparability, usability and interpretability of the input data and of the quality of the resulting statistical products. Based on the assessment results and after consulting with Member countries whereas appropriate, the Organization approves or not the use of the non-traditional data sources.

**Leadership and coordination**

All assessment activities related to new or existing processes and the use of non-traditional data sources are supported and coordinated by a team of experts that reports to the Data Coordination Group, Inter-Departmental Working Group (IDWG) on Data and Statistics and the Office of the Chief Statistician. The same team coordinates and promotes the implementation of improvement actions to ensure quality at institutional level, in line with priorities set by the IDWG. They also coordinate regular dialogues with Member States and the broader international data and statistics community on statistics and data quality-related issues and best quality assurance practices to address them.

**Implementation plan**

The implementation of the activities described above requires prior development and regular review of a set of tools, such as the checklist for self-assessment, methodological guidelines for the conduct of audits and an assessment tool and procedures for the use of non-traditional data sources. These activities are included in FAO programme of work with clear timelines and allocated resources, and their implementation are regularly monitored through the Data Coordination Group, IDWG and the reporting mechanisms of the FAO Programme of Work and Budget.
### Annex 1. Correspondence between FAO SDQAF principles and Principles Governing International Statistical Activities

<table>
<thead>
<tr>
<th>FAO SDQAF principles</th>
<th>Principles Governing International Statistical Activities (CCSA)(^{10})</th>
</tr>
</thead>
</table>
| **FAO SDQAF Principle 1:** Professional independence and impartiality | **Principle 1**  
High quality international statistics, which are accessible to all, are a fundamental element of global information systems.  
Good practices include:  
• Compiling and disseminating international statistics based on impartiality. |
| **FAO SDQAF Principle 2:** High and Continuous Commitment to Quality | **Principle 1**  
High quality international statistics, which are accessible to all, are a fundamental element of global information systems |
| **FAO SDQAF Principle 3:** Adequacy of resources | **Principle 6**  
Individual data collected about people 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 about people, households, businesses and other individual respondents. |
| **FAO SDQAF Principle 4:** Data protection and Statistical Confidentiality | **Principle 7**  
Erroneous interpretation and misuse of statistics are to be immediately and 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. |

---
\(^{10}\) [http://unstats.un.org/unsd/methods/statorg/Principles_stat_activities/principles_stat_activities.htm](http://unstats.un.org/unsd/methods/statorg/Principles_stat_activities/principles_stat_activities.htm)
<table>
<thead>
<tr>
<th>FAO SDQAF principles</th>
<th>Principles Governing International Statistical Activities (CCSA)</th>
</tr>
</thead>
</table>
| FAO SDQAF Principle 5: Development of international standards | Principle 5 Concepts, definitions, classifications, sources, methods and procedures employed in the production of international statistics meet professional scientific standards, and are made transparent for users. Good practices include:  
- Making officially agreed standards publicly available.  
Principle 8 Standards for national and international statistics are to be developed on the basis of sound professional criteria, while also addressing issues of practical utility and feasibility. Good practices include:  
- Systematically involving national statistical offices and other national organizations for official statistics in the development of international statistical programmes, including the development and promulgation of methods, standards and best 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.  
FAO SDQAF Principle 6: Cooperation with data providers | Principle 5 Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers. Good practices include:  
- Ensuring that national statistical offices and other national organizations for official statistics are duly involved, and that they advocate the application of the Fundamental Principles of Official Statistics when data are collected in countries.  
Principle 9 Coordination of international statistical programmes is essential to strengthen the quality, coherence and governance of international statistics, and to avoid duplication of work. Good practices include:  
- Coordinating technical cooperation activities in countries with resource partners and different organizations in the national statistical system, to avoid duplication of efforts and to encourage complementarities and synergies.  
Principle 10 Bilateral and multilateral cooperation in statistics contributes to the professional growth of the statisticians involved, and to the improvement of statistics within countries and the organizations involved. Good practices include:  
- Cooperating and sharing knowledge among international organizations and with countries and regions to further develop national and regional statistical systems.  
- Basing cooperation projects on user requirements, promoting the full participation of the main stakeholders, taking account of both local circumstances and the level of statistical development.  
- Empowering recipient national statistical systems and governments to take the lead. |
<table>
<thead>
<tr>
<th>FAO SDQAF principles</th>
<th>Principles Governing International Statistical Activities (CCSA)¹⁰</th>
</tr>
</thead>
</table>
| • Advocating the implementation of the Fundamental Principles of Official Statistics in countries.  
• Setting cooperation projects within a balanced overall strategic framework for the development of national official statistics. | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers.  
Good practices include:  
• Contributing to an integrated presentation of statistical programmes, including data collection plans, thereby making gaps or overlaps clearly visible  
Principle 9  
Coordination of international statistical programmes is essential to strengthen the quality, coherence and governance of international statistics, and to avoidduplication 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 organization and represents the organization in international statistical meetings.  
• Participating in international statistical meetings and bilateral and multilateral consultations whenever necessary.  
Principle 10  
Bilateral and multilateral cooperation in statistics contributes to the professional growth of the statisticians involved, and to the improvement of statistics within countries and the organizations involved.  
Good practices include:  
• Cooperating and sharing knowledge among international organizations and with countries and regions to further develop national and regional statistical systems. |
| FAO SDQAF Principle 7: Coordination with other international organizations producing statistics | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers.  
Good practices include:  
• Contributing to an integrated presentation of statistical programmes, including data collection plans, thereby making gaps or overlaps clearly visible  
Principle 9  
Coordination of international statistical programmes is essential to strengthen the quality, coherence and governance of international statistics, and to avoidduplication 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 organization and represents the organization in international statistical meetings.  
• Participating in international statistical meetings and bilateral and multilateral consultations whenever necessary.  
Principle 10  
Bilateral and multilateral cooperation in statistics contributes to the professional growth of the statisticians involved, and to the improvement of statistics within countries and the organizations involved.  
Good practices include:  
• Cooperating and sharing knowledge among international organizations and with countries and regions to further develop national and regional statistical systems. |
| FAO SDQAF Principle 8: Suitable and trustworthy data sources | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers.  
Good practice include  
• Having mechanisms in place to promote the use of the most suitable methods and sources by national statistical offices and other national organisations  
Principle 2  
To maintain the trust in international statistics, their production is to be impartial and strictly based on the highest professional standards  
Good Practice include  
• Using the best national data sources in compiling International Statistics, be they official or non-official sources, following the Recommended Practices on the Use of Non-Official Sources in International Statistics |
| FAO SDQAF Principle 9: Sound methodology and appropriate statistical procedures | Principle 4  
The concepts, definitions, classifications, sources, methods and procedures employed in the production of international statistics meet professional scientific standards, and are made transparent for the users.  
Good practices include:  
• Aiming to continuously introduce methodological improvements and systems to manage and improve the quality and transparency of statistics. |
<table>
<thead>
<tr>
<th>FAO SDQAF principles</th>
<th>Principles Governing International Statistical Activities (CCSA)¹⁰</th>
</tr>
</thead>
</table>
| • Enhancing the professional competencies of staff by encouraging them to attend training courses, to do analytical work, to publish scientific papers, and to participate in seminars and conferences.  
• In the dissemination of international statistics, giving credit to the original source, and using agreed quotation standards when re-using statistics originally collected by others | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers. |
| FAA SDQAF Principle 10: Cost-effectiveness | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers. |
| FAA SDQAF Principle 11: Non excessive burden on respondents | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers.  
Good practices include:  
• Facilitating the provision of data by countries.  
• Periodic review of statistical programmes to minimize the burden on data providers.  
• Sharing collected data with other organisations and collecting data jointly where appropriate |
| FAA SDQAF Principle 12: Relevance | Principle 1  
High quality international statistics, which are accessible to all, are a fundamental element of global information systems.  
Good practices include:  
• Regular consultations with both internal and external key users to ascertain that their needs are being met.  
• Periodic review of statistical programmes to ensure their relevance. |
| FAA SDQAF Principle 13: Accuracy and Reliability | Principle 4  
Concepts, definitions, classifications, sources, methods and procedures employed in the production of international statistics meet professional scientific standards, and are made transparent for users.  
Good practices include:  
• 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. |
| FAA SDQAF Principle 14: Timeliness and Punctuality | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers.  
Good practices include:  
• Working systematically on improving the timeliness of international statistics. |
| FAA SDQAF Principle 15: Coherence and Comparability | Principle 5  
Sources and methods for data collection are appropriately chosen to ensure timeliness and other aspects of quality, to be cost-efficient and to minimize the reporting burden for data providers.  
Good practices include:  
• Working systematically on improving the timeliness of international statistics. |
<table>
<thead>
<tr>
<th>FAO SDQAF principles</th>
<th>Principles Governing International Statistical Activities (CCSA)¹⁰</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Having mechanisms in place to consult countries to address discrepancies between national and international statistics</td>
<td></td>
</tr>
<tr>
<td>Principle 9</td>
<td>Coordination of international statistical programmes is essential to strengthen the quality, coherence and governance of international statistics, and to avoid duplication of work.</td>
</tr>
<tr>
<td>Good practices include:</td>
<td></td>
</tr>
<tr>
<td>• Working systematically towards agreements on common concepts, classifications, standards and methods. Working systematically towards agreement on which series should be considered as authoritative for each important set of statistics.</td>
<td></td>
</tr>
<tr>
<td>FAO SDQAF Principle 16: Accessibility and Clarity</td>
<td>Principle 1</td>
</tr>
<tr>
<td>High quality international statistics, which are accessible to all, are a fundamental element of global information systems.</td>
<td>Good practices include:</td>
</tr>
<tr>
<td>• Providing equal access to statistics for all users.</td>
<td>• Ensuring free public access to key statistics originally collected by others.</td>
</tr>
<tr>
<td>• Ensuring free public access to key statistics originally collected by others.</td>
<td>• Making officially agreed standards publicly available.</td>
</tr>
<tr>
<td>Principle 4</td>
<td>Concepts, definitions, classifications, sources, methods and procedures employed in the production of international statistics meet professional scientific standards, and are made transparent for users.</td>
</tr>
<tr>
<td>Good practices include:</td>
<td></td>
</tr>
<tr>
<td>• 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.</td>
<td>• In the dissemination of international statistics, giving credit to the original source, and using agreed quotation standards when re-using statistics</td>
</tr>
</tbody>
</table>
### Annex 2. Correspondence between FAO SDQAF principles and QAF principles of other international organisations

<table>
<thead>
<tr>
<th>FAO</th>
<th>ECB(^{11})</th>
<th>Eurostat(^{12})</th>
<th>IMF(^{13})</th>
<th>OECD(^{14})</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDQAF Principle 1: Professional independence and impartiality</td>
<td>SQF 1:</td>
<td>CoP 1:</td>
<td>DQAF 1.1: Assurance of Integrity - Institutional integrity (1.1.1, 1.1.2,</td>
<td>QF (a), (d)</td>
</tr>
<tr>
<td></td>
<td>Independence and accountability</td>
<td>Professional Independence</td>
<td>1.1.3) DQAF 1.2: Assurance of integrity - transparency (1.2.1 - 1.2.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SQF 3:</td>
<td>CoP 6:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Impartiality and objectivity</td>
<td>Impartiality and Objectivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SQF 10:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Accuracy and reliability (including stability) of the statistical output (4,5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQAF Principle 2: High and continuous commitment to quality</td>
<td>SQF 7:</td>
<td>CoP 4:</td>
<td>DAQF 0.4: Prerequisites for quality - Other quality management (0.4.1,</td>
<td>QF (g)</td>
</tr>
<tr>
<td></td>
<td>Sound methodology and appropriate statistical procedures (3)</td>
<td>Commitment to quality</td>
<td>0.4.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SQF 9:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relevance of the statistical output (4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQAF Principle 3: Adequacy of resources</td>
<td>SQF 6:</td>
<td>CoP 3:</td>
<td>DQAF 0.2: Prerequisites for quality - Resources</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Resources and efficiency</td>
<td>Adequacy of Resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQAF Principle 4: Data protection and statistical confidentiality</td>
<td>SQF 4:</td>
<td>CoP 5:</td>
<td>DQAF 0.1: Prerequisites for quality - Legal &amp; Institutional environment (0.1.3)</td>
<td>QF (c), (s)</td>
</tr>
<tr>
<td></td>
<td>Statistical confidentiality</td>
<td>Statistical confidentiality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQAF Principle 5: Development and use of international standards</td>
<td>SQF 7:</td>
<td>CoP 7:</td>
<td>DQAF 2: Methodological soundness (2.1, 2.2, 2.3)</td>
<td>QF (b), (n), (o)</td>
</tr>
<tr>
<td></td>
<td>Sound methodology and appropriate statistical procedures (1, 2, 3, 4)</td>
<td>Sound Methodology (7.1, 7.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SQF 11:</td>
<td>CoP 14:</td>
<td>Coherence &amp; comparability (14.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consistency (or coherence) and comparability of statistical outputs (5)</td>
<td>Coherence &amp; comparability (14.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SQF 10:</td>
<td>CoP 15:</td>
<td>Accessibility and Clarity (15.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consistency (or coherence) and comparability of statistical outputs (5)</td>
<td>Accessibility and Clarity (15.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

\(^{13}\) [https://dsbb.imf.org/content/pdfs/dqrs_Genframework.pdf](https://dsbb.imf.org/content/pdfs/dqrs_Genframework.pdf)  
\(^{14}\) [https://www.oecd.org/sdd/21687665.pdf](https://www.oecd.org/sdd/21687665.pdf)
| SDQAF Principle 6: Cooperation with data providers | SQF 1: Independence and accountability (2,3,4)  
SQF 2: Mandate for Data Collection  
SQF 5: Coordination and cooperation among the members of the ESCB and with European and international organizations  
SQF 6: Resources and efficiency (1,2) | CoP 2: Mandate for data collection  
CoP 3: Adequacy of resources (3.3-3.4) | DQAF 0.1: Prerequisites for quality – Legal & Institutional Environment (0.1.1-0.1.2) | QF (i) |
| SDQAF Principle 7: Cooperation with other international organizations producing statistics | SQF 5: Coordination and cooperation among the members of the ESCB and with European and international organizations | CoP 1bis: Coordination and cooperation (1bis.3) |  | QF (h), (i) |
| SDQAF Principle 8: Suitable and trustworthy data sources | SQF 3: Impartiality and objectivity (1, 2) | CoP 2: Mandate for data collection (2.1, 2.2, 2.4)  
CoP 7: Sound Methodology (7.1) | DQAF 3.1: Accuracy & Reliability - Source data  
DQAF 3.2: Accuracy & Reliability - Assessment of source data | QF (b), (n) |
| SDQAF Principle 9: Sound methodology and appropriate statistical procedures | SQF 7: Sound methodology and appropriate statistical procedures (1, 2, 3) | CoP 7: Sound methodology  
CoP 8: Appropriate Statistical procedures (8.2 - 8.6) | DQAF 2: Methodological soundness | QF (n) |
| SDQAF Principle 10: Cost-effectiveness | SQF 8: Cost-effectiveness and non-excessive burden of reporting agents | CoP 10: Cost effectiveness | DQAF 0.2: Prerequisites for quality - Resources (0.2.2) | QF (g), (j), (t), (u) |
| SDQAF Principle 11: Non excessive burden of respondents | SQF 8: Cost-effectiveness and non-excessive burden of reporting agents (3) | CoP 8: Appropriate Statistical procedures (8.7 - 8.9)  
CoP 9: Non-excessive burden on respondents. | DQAF. 0.1: Prerequisites for quality - Legal & Institutional environment (0.1.4) | QF (i), (l) |
| SDQAF Principle 12: Relevance | SQF 9: Relevance of the statistical output | CoP 11: Relevance | DQAF 0.3: Prerequisites for quality - Relevance (0.3.1) | QF (g), (l) |
| SDQAF Principle 13: Accuracy and Reliability | SQF 10: Accuracy and reliability (including stability) of the statistical output | CoP 8: Appropriate Statistical procedures (8.2 - 8.6) CoP 12: Accuracy & Reliability | DQAF 3.2: Accessibility & Reliability - Assessment of source data (3.2.1) DQAF 3.4: Accessibility & Reliability - Assessment and validation of intermediate data and statistical outputs DQAF 3.5: Accessibility & Reliability - Revision studies DQAF 4.3: Serviceability - Revision policy and practices | QF (g) |
| SDQAF Principle 14: Timeliness and Punctuality | SQF 12: Timeliness (including punctuality) of the statistical output | CoP 13: Timeliness & punctuality | DQAF 3.1: Accessibility & reliability - source data (3.1.3) DQAF 4.1: Serviceability - Periodicity and Timeliness DQAF 5.1: Accessibility - Data Accessibility (5.1.3) | QF (g) |
| SDQAF Principle 15: Coherence and Comparability | SQF 11: Consistency (or coherence) and comparability of statistical outputs | CoP 14: Coherence & comparability | DQAF 4.2: Serviceability - consistency | QF (g) |
| SDQAF Principle 16: Accessibility and Clarity | SQF 13: Accessibility and Clarity of the statistical output | CoP 15: Accessibility and Clarity | DQAF 5.1: Accessibility - Data accessibility DQAF 5.2: Accessibility - Metadata accessibility DQAF 5.3: Accessibility - Assistance to users (5.3.2) | QF (g), (o), (q), (r), (t), (u) |
In reviewing the 2014 FAO Statistics Quality Assurance Framework (SQAF) into the current FAO SDQAF, existing quality frameworks focusing on the use of innovative and non-statistical data sources were considered as references. Among these, the Big Data Quality Framework (BDQF) proposed by the UNECE Big Data Quality Task Team (UNECE, 2014) played a prominent role in the adaptation of FAO’s statistical quality principles. Indeed, while preserving the structure of the original FAO’s SQAF, most of its principles have been expanded in their scope, to better align with the quality dimensions envisaged in the UNECE BDQF. The present annex provides a summary of the main correspondences between the new FAO SDQAF and the UNECE BDQF.

The BDQF presents a structured view of quality around the three main phases of the statistical production process, namely:

1) **The input stage**, considering all decisions and actions for the acquisition, or pre-acquisition analysis of input big data source.

2) **The throughput stage**, dealing with data processing, transformation, and analysis.

3) **The output stage**, concerned with the production and dissemination of statistical outputs.

It should be noted that, as stated in the introduction of this document, the FAO SDQAF provides an overall framework to ensure the quality of FAO statistical outputs, and the soundness of the processes and governance mechanisms put in place for their production. On the other hand, the quality elements specific to different typologies of input data used by FAO are addressed in dedicated and complementary corporate statistical standards. These elements are summarized in the FAO SDQAF under Principle 7 on Suitable and trustworthy data sources (see Table A.3), which—with its key implementation modalities—ensures that all the quality dimensions considered in the UNECE BDQF are well addressed. On the other hand, a more detailed mapping between the quality principles and dimensions laid out for the throughput and output stage in the BDQF and principles in the FAO SDQAF was possible, as presented in Table A.3.

The BDQF uses a hierarchical structure composed of three hyperdimensions with nested quality dimensions. The three hyperdimensions, which have been borrowed from the administrative data quality framework produced by Statistics Netherlands, are the source, the metadata and the data.

Irrespective from the phase of the business process or the considered hyperdimension, the BDQF is governed by three overarching principles, which are well reflected in the current FAO SDQAF:

1. **Fitness for use**, i.e. whether the data source is appropriate for the intended purpose. This principle is also at the core of the FAO SDQAF, which—in line with ISO’s practices
defines quality as “the totality of features and characteristics of a product or service that bears to satisfy stated or implied needs” (ISO No 8402; 1986, 3.1).

2. **Generic and flexible**, meaning that a quality framework should be broad enough to be applicable over a wide variety of situations. This principle is well embraced by the FAO SDQAF that, instead of being prescriptive, outlines 16 broad and flexible principles that well adapt to different data sources and scenarios, and are then put into best practices through specific corporate statistical standards.

3. **Efforts versus gain**, i.e., whether the effort involved in obtaining and analysing the big data source worth the benefits gained from the data source itself. This principle is well reflected in the FAO SDQAF Principle 7 envisaging, in its point 7.3, that cost-benefit and risk analysis is carried out before new, or substantially enhanced, statistical processes are integrated into regular production
Table A.3: Mapping of BDQF quality dimensions and FAO SDQAF principles

<table>
<thead>
<tr>
<th>Phase</th>
<th>Hyperdimension</th>
<th>Quality Dimension</th>
<th>Description</th>
<th>Link with FAO SDQAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Source</td>
<td>Institutional/Business environment</td>
<td>At the input stage, the quality of the institutional/business environment of the organization providing the data is considered. Elements to be assessed are: 1) Sustainability of the entity/data provider; 2) Reliability Status; 3) Transparency and interpretability.</td>
<td>SDQAF Principle 8: Suitable and trustworthy data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Privacy and Security</td>
<td>This is a prominent quality dimension when considering big data sources. A key aspect to be considered at the input stage is that of consent for data acquisition, and whether the consent is active or passive. Consent may be obtained via an &quot;agreement to terms&quot; or in more explicit ways. An organization acquiring big data should pay attention to whether this consent is given and whether this is in line with its regulations. Elements to consider are: 1) Legislation related to data production, maintenance, access, and secondary uses; 2) Restrictions in terms of privacy/confidentiality that would limit data use; 3) Perception of intended use.</td>
<td>SDQAF Principle 8: Suitable and trustworthy data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complexity</td>
<td>Refers to the lack of simplicity and uniformity in the data. Complexity at the input stage can be assessed looking at: 1) Technical constraints related to receiving, reading, processing and storing the files; 2) Data structure and format, referring to the complexity of the structure and format in which data are stored; 3) Readability of data; 4) Hierarchies and nesting.</td>
<td>SDQAF Principle 8: Suitable and trustworthy data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>The extent to which metadata are available, interpretable and complete for proper understanding and use of data. Aspects to be clarified in the metadata are: 1) Processes that led to the collection of the data; 2) Processes related to data treatment; 3) Description of the data itself.</td>
<td>Specific best practices to fulfil each quality dimensions are provided in the Corporate Statistical Standard on the use of Big Data sources to produce statistical outputs.</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Phase</th>
<th>Hyperdimension</th>
<th>Quality Dimension</th>
<th>Description</th>
<th>Link with FAO SDQAF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Usability</td>
<td>The extent to which the receiving organization will be able to work with and use the data. The assessment of this dimension entails: 1) Evaluating the additional resources needed; 2) Assessing the risk of having to make considerable investments (also in terms of skills) in order to use the big data.</td>
<td>SDQAF Principle 3: Adequacy of resources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time-related factors</td>
<td>Time-related factors include: 1) Timeliness; 2) Punctuality; 3) Periodicity; 4) Changes over time.</td>
<td>SDQAF Principle 14: Timeliness and Punctuality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coherence - Linkability</td>
<td>Ease with which a big data source can be linked, when relevant, with other datasets. Aspects to be considered are the presence and the quality of linking variables at the necessary levels.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coherence - Consistency</td>
<td>The extent to which a dataset complies with standard definitions and is consistent over time. At the input stage, the use of standardized concepts is mainly related to the use of standards for key variables.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validity</td>
<td>The extent to which a dataset measures what the user is attempting to measure. This concept, which in standard quality frameworks is often included under the coherence dimension, is particularly relevant for big data. Validity can be assessed in terms of: 1) Transparency of methods and processes; and 2) Soundness of adopted methods.</td>
<td>SDQAF Principle 8: Suitable and trustworthy data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accuracy and selectivity</td>
<td>Degree to which the information correctly describes the phenomena it was designed to measure. When assessing the accuracy of a potential dataset to be used for statistical purposes, it is recommended to use a &quot;total survey error&quot; approach. At the input stage, this translates in the consideration of the following elements: over-coverage, under-coverage, selectivity, missing data (non-observation and/or non-response), data adjustments and presence of anomalies.</td>
<td>Specific best practices to fulfill each quality dimensions are provided in the Corporate Statistical Standard on the use of Big Data sources to produce statistical outputs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coherence - Linkability</td>
<td>Quality of linking variables</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coherence - Consistency</td>
<td>Coherence of the metadata with the observed data</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validity</td>
<td>Coherence between process and methods described in the metadata and the observed data values.</td>
<td></td>
</tr>
<tr>
<td>Phase</td>
<td>Hyperdimension</td>
<td>Quality Dimension</td>
<td>Description</td>
<td>Link with FAO SDQAF</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Throughput</td>
<td>NA</td>
<td>System independence</td>
<td>Transformation and analysis should depend on the adopted theoretical principles and not the specific software/hardware used.</td>
<td>SDQAF Principle 2: High and Continuous Commitment to Quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Steady states</td>
<td>A steady state is a version of a dataset that has met certain quality criteria. Datasets can always be processed, analyzed, transformed, etc., but they should transition from a steady state to another with proper versioning.</td>
<td>SDQAF Principle 8: Suitable and trustworthy data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality Gates</td>
<td>Quality gates are check points in the business process cycle at which the quality of data is explicitly assessed. Measures and features of quality gates as well as the position of quality gates need to be set in advance.</td>
<td>SDQAF Principle 9: Sound methodology and appropriate statistical procedures</td>
</tr>
<tr>
<td>Output</td>
<td>Institutional/Business environment</td>
<td>Privacy and Security</td>
<td>At the output phase, the quality of the institutional environment of the institution disseminating the statistical output is considered. Elements to be looked at are: 1) The nature of the input big data source (e.g., social media, satellite data, etc.); 2) The arrangements under which data was transferred to the disseminating institution; 3) Quality assurance processes applied to the incoming data; 4) Role that the data played in the final output product.</td>
<td>SDQAF Principle 6: Cooperation with data providers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Complexity</td>
<td>This dimension should consider: 1) The legislation related to data production, maintenance, and access; 2) Restrictions limiting data use (privacy, security, confidentiality); 3) Actions taken to mitigate potential negative perceptions on the use of data from stakeholders.</td>
<td>SDQAF Principle 8: Suitable and trustworthy data sources</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Refers to the lack of simplicity and uniformity in the data. At the output phase, the following elements should be considered: 1) Data treatment, i.e., how the complexity of the input data has been dealt with during the input and throughput stages regarding data structure, format, data hierarchies; 2) Actual limitations to the use of statistical outputs caused by the complexity of big data used as input.</td>
<td>SDQAF Principle 4: Data protection statistical confidentiality</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SDQAF Principle 9: Sound methodology and appropriate statistical procedures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SDQAF Principle 13: Accuracy and Reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SDQAF Principle 16: Accessibility and Clarity</td>
</tr>
<tr>
<td>Phase</td>
<td>Hyperdimension</td>
<td>Quality Dimension</td>
<td>Description</td>
<td>Link with FAO SDQAF</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------</td>
<td>-------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Data</td>
<td></td>
<td>Accessibility and clarity</td>
<td>Factors to be considered are the extent to which: 1) Accessibility of data and metadata; 2) Presence of clear, unambiguous definitions, explanations and quality indicators; 3) Conformity to metadata standards.</td>
<td>SDQAF Principle 16: Accessibility and Clarity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relevance</td>
<td>Whether the data measures the concepts meant to be measured for the intended use.</td>
<td>SDQAF Principle 12: Relevance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accuracy and selectivity</td>
<td>At this stage, traditional measures of statistical accuracy should be considered, such as standard error, bias, etc. In addition, selectivity issues should be considered when big-data are used as input to produce statistical outputs.</td>
<td>SDQAF Principle 13: Accuracy and Reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Validity</td>
<td>At the output stage, the following elements should be considered: 1) Convergent validity, i.e., how well the metric aligns with other similar metrics; 2) Conceptual utility, i.e., the extent to which the metric is able to provide insight into real-world phenomena. 3) Methodological validity, i.e., the extent to which the methods underlying the metric are transparent and theoretically sound.</td>
<td>SDQAF Principle 12: Relevance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coherence - Linkability</td>
<td>Ease with which a statistical output can be linked, when relevant, with other databases</td>
<td>SDQAF Principle 15: Coherence and Comparability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coherence - Consistency</td>
<td>The extent to which a statistical output complies with standard definitions and is consistent over time.</td>
<td>SDQAF Principle 15: Coherence and Comparability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time-related factors</td>
<td>Time-related factors include: 1) Timeliness; 2) Punctuality; 3) Periodicity; 4) Changes over time.</td>
<td>SDQAF Principle 14: Timeliness and Punctuality</td>
</tr>
</tbody>
</table>
Annex 4. Glossary

Administrative data sources: data sets created primarily for administrative purposes by government agencies or other entities working on behalf of the government. Administrative data sources include administrative registers of persons and legal entities and the records of ministries, departments and specialized agencies, such as tax returns, social services records and customs data, or data of regional or local administrations. In contrast to statistical data sources, administrative data sources are not created in response to the need for statistical data but as a part of a government function, such as the provision of services or taxation. In some cases, statistical agencies participate in the design and/or collection of administrative data. In addition, statistical agencies may be involved at different stages of the production process of administrative data, with the aim of ensuring that the data will be usable for statistical purposes;

Big data: data generated by business or government, transactions, social media, phone logs, communication devices, web crawling, sensors, etc. characterised by high volume, velocity and variety (UN Handbook on Management and Organisation of National Statistical Systems (2021))\(^\text{15}\). Sources described as ‘Big Data’ can be largely unstructured, meaning that they have no pre-defined data model and do not fit well into conventional relational databases. In addition, they often require new tools and methods to capture, curate, manage and process them in an efficient way\(^\text{16}\). The UNECE Task team on big data (2013)\(^\text{17}\) proposes a classification of big data sources based on how there are generated:

- **human-sourced** information available mostly from social networks, blogs, internet searches, etc. where data are loosely structured and often un governed;
- **process-mediated** data available from the IT systems of organisations (private or public), where data is usually structured and stored in relational databases;
- **machine-generated** data captured by sensors and other machines used to measure and record events in the physical world.

Data: characteristics or information, usually numerical, that are collected through observation (OECD Glossary of Statistical Terms)\(^\text{18}\). Indeed, these characteristics or information can be generated by traditional statistical surveys or derived from other sources. In the context of the SQDAF, data is used when referring to data sources (input phase) or during the statistics production process (throughput phase) (the term data includes microdata which, depending on the context, can be also a statistical output)

Data confidentiality: Property of data indicating whether they are subject to dissemination restrictions. Unauthorised disclosure of this type of data could be prejudicial or harmful to the interest of the source or other relevant parties. For instance, data allowing the identification of a physical or legal person, either directly or indirectly, may be characterised as confidential according to the relevant national or international legislation. Unauthorised disclosure of data that are restricted or confidential is not permitted and even legislative measures or other formal provisions may be used to prevent disclosure. Often, there are procedures in place to prevent disclosure of restricted or confidential data, including rules applying to staff, aggregation rules when disseminating data, provision of unit records, etc.

Data source: a specific data set, metadata set, database or metadata repository from where data or metadata are available. They can take the form of statistical survey, administrative sources or other

---


\(^\text{17}\) [https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data](https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data)

data sources, such as big data. For the purpose of this framework, two types of data sources are considered:

- **Statistical Data Sources**: data collections created primarily for official statistical purposes by government agencies or other entities working on behalf of the government. Statistical data sources include statistical sample surveys, censuses and statistical registers. There are different types of censuses, such as population and housing censuses, business censuses and agriculture censuses, among other types. Sample surveys and statistical registers can cover different units, for example individuals, households and businesses. Statistical registers can themselves be derived from different sources. For example, statistical business registers are often based on administrative data sources;

- **Non-statistical data sources (or non-traditional data sources)**: all data sets that are not created primarily for official statistical purposes but rather for administrative, commercial or other private purposes. In general, non-statistical data sources include administrative data sources and big data that have not yet being processed and transformed into a statistical outputs.

**Metadata**: Data that defines and describes other data and processes. There are two types of metadata that can be distinguished from each other:

- **Structural metadata** define and accompany the data and consist of identifiers and descriptors that are essential for discovering, organizing, retrieving and processing a statistical data set (e.g., titles, subtitles, short descriptions, dimension names, variable names, etc.)

- **Reference metadata** are of a more general nature and describe statistical concepts and methodologies used for the collection and generation of data and provide information on data quality, thereby assisting users with the interpretation of the data. Contrary to structural metadata, reference metadata can be decoupled from the data (i.e., they can be generated, collected or disseminated separately from the statistics to which they refer).

**Non-official data**: Data that were not officially reported or published by the national statistical organization (or any agency responsible for statistics)\(^\text{19}\). They are rather produced by non-public national sources including private sector, NGOs, academic research or data that are originally constructed by International Organizations using different methods (imputation, estimation, construction of new indicators, etc.)\(^\text{20}\).

**Non-traditional data sources**: see non-statistical data sources.

**Official data**: In the context of international organizations, an official data refers to a set of national official statistics produced by recognized government agencies, members of the National Statistical System, in compliance with the Fundamental Principles of Official Statistics and national quality frameworks, as well as other internationally agreed statistical standards and recommendations. In many countries, official statistics are defined and described in the statistical programmes and collected under a statistics legislation. Official data sources include data disseminated as official statistics by the national statistical agencies or official national focal points or transmitted to FAO as such (e.g. by other recognized international organizations). They can be the result of a statistical process based on statistical or non-statistical data sources or a joint data collection by national and international recognized agencies. When disseminating its statistics, FAO distinguish data points originating from official sources from those produced from non-official data.

**Quality**: It is a multi-faceted and subjective concept. The International Organisation for Standardization (ISO) defines quality as “the totality of features and characteristics of a product or service that bears to satisfy stated or implied needs” (ISO No 8402; 1986, 3.1). The most important quality characteristics depend on user perspectives, needs and priorities, which vary across groups


of users (‘fitness for purpose’). In the context of the SDQAF, the quality dimensions taken into consideration when assessing the quality of a statistical output are: Relevance, Accuracy and Reliability, Timeliness and Punctuality, Coherence and Comparability, Accessibility and Clarity. FAO’s Quality Assurance process aims at achieving the best possible compromise in satisfying the needs of the various possible users (current and potential) with regards to these quality dimensions given constraints concerning available resources.

**Quality dimensions:** for statistics, the general definition of quality is operationalized by specifying a set of factors or dimensions that characterize the quality of the product. The SDQAF identifies quality dimensions linked to statistical products in quality principles 11 to 15 covering the following dimensions (some principles cover two closely related dimensions):

- **Relevance:** Quality dimension that measure the degree to which statistics meet the current and potential user needs.
- **Accuracy:** Quality dimension that refers to the closeness of estimates, to the true values that statistics were intended to measure.
- **Reliability:** Quality dimension that refers to the closeness of the initial estimates to the subsequent or final estimates.
- **Timeliness:** Quality dimension that refers to the speed of dissemination of statistical outputs - i.e. the lapse of time between the end of a reference period (or a reference date) and the dissemination of the statistical outputs.
- **Punctuality:** Quality dimension that refers to the possible time lag existing between the actual delivery date of statistical outputs and the target date when they should have been delivered, for instance, with reference to dates announced in an official release calendar or previously agreed among partners.
- **Coherence:** Quality dimension that corresponds to the adequacy of the statistical outputs to be meaningfully combined in different ways and for various uses.
- **Comparability:** Quality dimension that refers to the extent to which differences between different geographical areas, non-geographical domains, or over time, can be dimensioned to differences between the true values of the statistical characteristics.
- **Accessibility:** Quality dimension that is defined as the ease, the set of conditions and the modalities by which users can obtain data.
- **Clarity:** Quality dimension that refers to the availability of adequate documentation: whether data are accompanied with appropriate metadata, illustrations such as graphs and maps, whether information on their quality are also available (including limitation in use), and the extent to which additional assistance is provided.

**Statistical output:** see statistics.

**Statistical process:** set of procedures used during the production of statistics, including the specification of needs, design, building, data collection, processing, analysis, dissemination of the products and evaluation of the process.

**Statistics (or Statistical output):** Numerical information relating to an aggregate of data on units or observations. In general, the SDQAF uses the term statistics or statistical outputs when referring to an output of a statistics production process. While data and statistics are sometimes used interchangeably in the literature, the later are closely related to the concepts of representativeness and aggregation at the output stage while the term data is used when referring to data sources or during the statistics production process (the term data includes microdata which, depending on the context, can be also a statistical output).