Guidelines and Best Practices on Data Flows and Global Data Reporting for Sustainable Development Goals

A. Introduction

I. Aim and scope of the guidelines

The Millennium Development Goals (MDG) process shed light on the importance of robust and reliable data for evidence-based decision-making as well as for effectively focusing national development policies and programs. Through the monitoring requirements many countries furthermore strengthened their statistical capacities and increased the availability of more and better data. But the MDG process also showed that the key concepts of cooperation, coordination and transparency between international organisations and National Statistical Systems (NSS) are of utmost importance in order to provide reliable, high-quality and impartial data for decision-makers. In consequence for the Sustainable Development Goals (SDG) process it is crucial to understand the data flows and reporting structures between national and international organisations and to find a common ground on how to work together in order to have harmonised and comparable statistics and indicators at the national, regional and international level.

The UN Statistical Commission requested the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDG) in its resolution 48/101 (l) “to develop detailed guidelines of how custodian agencies and countries can work together to contribute to the data flows necessary to have harmonized statistics” for global SDG data reporting.

While the UN resolution A/RES/70/1 clearly states that “national ownership is key to achieving sustainable development” and therefore “the global review will be primarily based on national official data sources”, other data sources might be necessary in order to close gaps or even validate and adjust existing data. The purpose of this document is to establish principles for global SDG data reporting and identify an approach for National and International Statistical Systems to ensure the quality of the official data and official statistics used for global reporting, as well as to decide upon the use of non-official data based on its quality. Moreover the guidelines will build on the extensive knowledge gained from past experiences and address coordination as well as cooperation between National and International Statistical Systems in this regard.

In order to ensure a high quality in international data series it is crucial that international data be based on comparable and consistent reported national data. Differences between national and international statistical data (caused by e.g. different methodologies or definitions, adjustments, estimates, missing values, invalid or inconsistent entries, potential errors or

other methodological as well as conceptual issues) do not only cause confusion, but can also damage the credibility of statistical agencies. They also can result in major discrepancies between organisations. Therefore the focus should be on a consistent and coordinated approach for processing data and metadata particularly with regards to validation and editing. This means clearly elaborated data flows are necessary to enhance coordination between National and International Statistical Systems. Equally important is the necessity to identify ways to deal with and communicate adjustments and estimations undertaken by organisations reporting on SDGs.

II. Provisions

The adoption of the General Assembly (GA) resolution A/RES/70/1 “Transforming our world: the 2030 Agenda for Sustainable Development” was a bold step and demonstrates the necessity to shift the world towards a sustainable development, in all its dimensions. With this resolution the international community also committed to a follow-up and review process at the national, regional and international level in order to monitor the progress made in implementing the Sustainable Development Goals and targets over the coming years.

In its resolution A/RES/71/313 the GA instructed the UN Statistical Commission (StatCom) with several tasks. Some of the tasks given to the StatCom include, “to coordinate the substantive and technical work to develop international statistical standards, methods and guidelines, where necessary […]” and “through the Inter-Agency and Expert Group on Sustainable Development Goal Indicators, to further refine and improve the global indicator framework in order to address coverage, alignment with targets, definition of terms and development of metadata and to facilitate its implementation […]” (A/RES/71/313 para. 3).

Since taking ownership is key for this transformation at all levels, member states play the central role in the Agenda 2030 process. This also means that individual national realities, capacities and policy priorities with regards to the follow-up and review process will have to be respected (A/RES/70/1 para. 74). At the same time the GA emphasises the importance of a strong and revitalised global partnership, which “brings together Governments, the private sector, civil society, the United Nations system and other actors” (A/RES/70/1 para. 39).

As the global review is to be primarily based on national official data sources, the GA stresses that National Statistical Offices (NSO) hold a coordinating role in the National Statistical Systems (A/RES/71/313 para. 6).

A critical aspect for the achievement of the Agenda 2030 goals is to make high-quality official statistical information available. Therefore, all levels and stakeholders that are part of the global statistical system have to adhere to the Fundamental Principles of Official Statistics, endorsed by the GA in its resolution A/RES/68/261 of 29 January 2014. This was again stressed in resolution A/RES/71/313, stating “that all activities of the global statistical system must be conducted in full adherence to the Fundamental Principles of Official Statistics” (A/RES/71/313 para. 8).

In this context the Generic National Quality Assurance Framework template, which was fully endorsed by the StatCom in its forty-third session in 2012, gives guidance on ensuring these quality demands. With respect to the National Quality Assurance Frameworks template and guidelines, the Expert Group on National Quality Assurance Framework was assigned the task to update the template and guidelines.

Regarding the production of high-quality data the GA also recalled the Economic and Social Council (ECOSOC) resolution 2006/6. The ECOSOC states that imputation by all international agencies should be avoided “unless specific country data are available for reliable imputations following consultations with concerned countries and through transparent methodologies” (ECOSOC Resolution 2006/6 para. 5c).

The monitoring process for the MDGs illustrated that the differences between nationally reported data and international disseminated data “can be considerably large” (UNECE; 2011). Reasons behind these differences include the use of different primary data sources, the operations of data revision and adjustments or different time coverages by international and national agencies. For the SDGs the GA therefore put on record that the follow-up and review process should primarily be based on data produced by National Statistical Systems. In cases where specific country data are not available, international agencies are urged to consult with the concerned countries before publishing own estimations, in order to make harmonised data available (A/RES/71/313 para. 7). The StatCom also recommendes that in cases where country data are adjusted or estimated, this should be done in full consultation with the concerned country (StatCom 48/101 para. l). The involvement of the concerned country is also highlighted in the report of the Secretary-General on “Quality assurance in the global statistical system” to the 48th Session of the StatCom.

Furthermore, international agencies are urged to enhance communication and coordination amongst each other “in order to avoid duplicate reports, ensure consistency of data and reduce response burdens on countries” as well as “provide the methodologies used to harmonise country data for international comparability and produce estimates through transparent mechanisms” (A/RES/71/313 para. 7). This necessity is also reinforced in a statement by the Secretary-General (E/CN.3/2017/8).

The StatCom also decided that when other sources and methodologies are used these will be reviewed and agreed by national statistical authorities and presented in a transparent manner (48/101 (l)).

The StatCom further requested the custodian agencies to provide “a list of national agencies providing the data to the international system and share data collection calendars in order to ensure the full traceability of data used in international sources” (48/101 (k)).

Finally, the report “A world that counts” elaborates nine key principles for the data revolution for sustainable development. As a part of these, the report underlines the necessity of periodic audits by professional and independent third parties in order to demonstrate the high quality and integrity of statistical information, especially in the light of non-traditional data sources such as big and geospatial data (Independent Expert Advisory Group on a Data Revolution for Sustainable Development, 2014).

B. Guidelines for Global SDG Data Reporting

I. Principles to guide data reporting from National to International Statistical Systems and data sharing for SDG indicators

1. Definitions:

    a) ‘official data’ refers to a set of values of qualitative or quantitative variables, which are produced and / or disseminated by an official source such as the National Statistical Office or another governmental agency or department including non-traditional types of data.

    b) ‘official statistics’ means a part of official data, which is produced in compliance with Fundamental Principles of Official Statistics.

2. All organisations that are involved in the global follow-up and review process of SDGs, including national statistics offices, other government departments, and international organizations are committed to:

a) Basing the global review primarily on national official data,

b) Acknowledging the coordinating role of National Statistical Offices in the National Statistical Systems,

c) Promoting the use of an appropriate data transmission standard, such as Statistical Data and Metadata eXchange (SDMX), for the collection and exchange of SDG data,

d) Providing metadata according to scientific standards on the sources, methods and procedures of the statistics,

e) Using national SDG indicator reporting platforms where available and take into account already well-established data reporting mechanisms. While National Statistical Systems are committed to timely dissemination of national official data, via a national reporting platform if available, custodian agencies are committed to retrieving the official data from those same reporting platforms. If no official data is available in the national reporting platform and there is no other well-established data reporting mechanism, the custodian agency will check with the National Statistical Office to determine official data availability,

f) If no national SDG indicator reporting platform is available and there is no other well-established data reporting mechanism, the National Statistical Office and the custodian agency determine together a data flow model. Countries may choose to declare regional database as primary source for custodian agencies to retrieve country data,

g) Establishing a common timetable for data reporting of National Statistical Systems, including timeframes for technical consultations and accreditation procedures, while coordinating and arranging data requests and technical consultations in a workable timely manner,

h) Determining one custodian agency responsible for each SDG indicator and a contact point in the corresponding agency. This list will be disclosed to the National Statistical Office [by UNSD] [by the end of 2017],

i) Contacting and communicating with the National Statistical Systems via national focal points preferably from the National Statistical Office taking into account already well-established data reporting mechanisms. In case no focal point is identified, the National Statistical Office will be considered the focal

point. Furthermore the NSO is either copied in all SDG data requests or provided with a list of all national data providers by custodian agencies,

j) In cases, where the SDG indicators are used by custodian agencies as well as partner agencies in thematically connected publications and databases the identical data sets from the global SDG database provided by the National Statistical System are to be used,

k) Coordinating their data collection work and establish effective and efficient data sharing arrangements among custodian agencies to avoid duplication of efforts. Data provided to the custodian agencies by the NSS shall only be shared after quality assurance and the plausibility check are carried out and finished in order to ensure consistency of data,

l) Coordinating updates on UN databases to ensure consistency of data.

3. A guiding principle for National and International Statistical Systems which are involved with the global SDGs data flow and reporting is to ensure a high quality of statistical information.

4. The high quality of official data disseminated in international statistics shall be ensured by:

a) Promoting the use of international standards for data in accordance with the UN Classifications Registry or other standards adopted by the UN Statistical Commission, selecting the most appropriate statistical methods and data sources exclusively on the basis of scientific and statistically robust considerations as well as a transparent documentation, and

b) Having in place:

(i) laws or guidelines applicable in the area of quality assurance,

(ii) monitoring systems,

(iii) enforcement systems, and

c) A mandatory quality assessment of the data set carried out by the responsible custodian agency and consisting of a set of tools spelled out in the Guidelines for the Template for a Generic National Quality Assurance Framework. Based on this quality assessment the decision about the use of a data set for global SDG data reporting shall be:

(i) in case of official data, in consensus with the national official data provider as well as the National Statistical Office, and

(ii) in case of official statistics consensual with the National Statistical Office.

In case of non-response from the national official data provider or the National Statistical Office, the custodian agency may understand this as if the corresponding entity had expressed their consent. This only applies if the contact information used to send the request had been previously validated, and there is no response following two subsequent requests with reasonable timeframes.

5. The quality of non-official data disseminated in international statistics shall be ensured by:

a) Promoting the use of international standards for data in accordance with the UN Classifications Registry, selecting the most appropriate statistical methods and data sources exclusively on the basis of scientific and statistically robust considerations as well as a transparent documentation, and

b) Having in place:

(i) laws or guidelines applicable in the area of quality assurance,

(ii) monitoring systems,

(iii) enforcement systems, and

c) An accreditation procedure[1] for non-official data by custodian agencies which includes a consultation with all involved parties as specified in Annex 1. The National Statistical Office may carry out the accreditation procedure if they wish to. The decision whether to utilise the data set and for which purposes shall be based on this accreditation procedure and taken in consensus with the listed national focal point and the National Statistical Office. The National Statistical Office should also be involved during the different steps of the accreditation procedure.

d) In cases where specific national official data is not available and custodian agencies or other international agencies estimate or model country data the

[1] Estimates or modelled data from non-official sources will be substituted, when / if national official data becomes available and following the mandatory quality assessment as outlined in Article 4 para. c.

same accreditation procedure applies as for non-official data. The decision whether to utilise the estimates or modelled data and for which purposes shall be taken in consensus with the listed national focal point and the National Statistical Office.

e) In case of non-response from the national focal point or the National Statistical Office to the accreditation procedure, the custodian agency may understand this as if the corresponding entity had expressed their consent. This only applies if the contact information used to send the request had been previously validated, and there is no response following two subsequent requests with reasonable timeframes.

II. Principles to guide the production of international data series for SDG indicators

(1) Concerning Data

6. When a plausibility check such as specified in Annex 2 identifies missing values or invalid or inconsistent entries or potential errors or data gaps and makes it necessary for the custodian agency to adjust, including through imputation, these country data the following procedure[2] shall apply:

a) The custodian agency generates a list of potential errors as a starting point for the adjustment process.

b) The custodian agency selects the most appropriate statistical methods and data sources exclusively on the basis of scientific and statistically robust considerations as for example referred to in Practical Guide on Data Validation and Editing, in order to calculate the adjusted values. The adjusted values are added to the list of potential errors, in order to compare them. The list of potential errors, the adjusted values and the chosen methodological approach shall be disclosed to the National Statistical Office and the national official data provider. Furthermore the chosen methodological approach shall be disclosed to anyone upon request.

c) In case of official data, the custodian agency invites the listed national official data provider to participate in a joint technical consultation via conference call,

[2] This procedure does not apply to standard practices of conversion to common measurement units.

electronic means or physical meetings with the aim to improve and validate the adjusted values as well as the used method.

d) In case of official statistics, the custodian agency invites the National Statistical Office to participate in a joint technical consultation via conference call, electronic means or physical meetings with the aim to improve and validate the adjusted values as well as the used method.

e) In case of non-official data[3], the custodian agency invites the listed national focal point, the National Statistical Office as well as the data set provider to participate in a joint technical consultation via conference call, electronic means or physical meetings with the aim to improve and validate the adjusted values as well as the used method.

f) The custodian agency ensures a reasonable timeframe for the National Statistical System and data providers to respond, but at least two weeks minimum. In case of non-response from the National Statistical System or data providers, the custodian agency may understand this as if the corresponding entity had expressed their consent. This only applies if the contact information used to send the request had been previously validated, and there is no response following two subsequent requests with reasonable timeframes.

g) The solution agreed upon will be adopted and published in the global SDG database and all thematically connected international publications in order to avoid discrepancies.

7. If in the case of missing values or invalid or inconsistent entries or potential errors or data gaps no agreement is found during the technical consultation according to Article 6 para. c - e the country data requiring adjustment shall be:

    (i) disseminated alongside the estimates from the custodian agency accompanied by an explanation of the reasons for the discrepancy and the chosen methodologies by the concerned parties.
    (ii) If both parties wish so, the concerned United Nations regional commission can be called upon to reach a regional solution on the subject matter.

[3] According to Article 5 para. c the decision to utilise non-official data shall be taken in consensus with the national official focal point and the NSO.

8. When during a plausibility check, such as specified in Annex 2, issues such as methodological gaps as well as conceptual issues and differences in the metadata as for example between nationally and internationally used standards or because of changes of standards over time are found and these make it necessary for the custodian agency to adjust data, the following procedure shall be undertaken:

    a) A list is to be generated on these differences as starting point for the adjustment process of the data.

    b) The custodian agency selects the most appropriate statistical methods exclusively on the basis of scientific and statistically robust considerations as for example referred to in Practical Guide on Data Validation and Editing, in order to calculate the adjusted values. The adjusted values are added to the list of differences, in order to compare them. Besides the list of differences and the adjusted values, the chosen methodological approach shall be disclosed to the National Statistical Office, the national official data provider and anyone upon request.

    c) In case of official data, the custodian agency invites the listed national official data provider to participate in a joint technical consultation via conference call, electronic means or physical meetings with the aim to validate the adjusted values as well as the used method.

    d) In case of official statistics, the custodian agency invites the National Statistical Office to participate in a joint technical consultation via conference call, electronic means or physical meetings with the aim to validate the adjusted values as well as the used method.

    e) In case of non-official data[4], the custodian agency invites the listed national focal point, the National Statistical Office as well as the data set provider to participate in a joint technical consultation via conference call, electronic means or physical meetings with the aim to validate the adjusted values as well as the used method.

    f) The custodian agency ensures a reasonable timeframe for the National Statistical System and data providers to respond, but at least two weeks minimum. In case of non-response from the National Statistical System or data providers, the custodian agency may understand this as if the

[4] According to Article 5 para. c the decision to utilise non-official data shall be taken in consensus with the national official focal point and the NSO.

corresponding entity had expressed their consent. This only applies if the contact information used to send the request had been previously validated, and there is no response following two subsequent requests with reasonable timeframes.

g) The solution agreed upon will be adopted and published in the global SDG database and all thematically connected international publications, in order to avoid discrepancies.

9. If no agreement is found during the technical consultation according to Article 8 para. a-g the country data requiring adjustment shall be:

(i) disseminated alongside the adjusted data from the custodian agency accompanied by an explanation of the reasons for the discrepancy and the chosen methodologies by the concerned parties.
(ii) If both parties wish so, the concerned United Nations regional commission will be called upon to reach a regional solution on the subject matter.

(2) Concerning Metadata

10. For the National Statistical System to provide accurate and internationally comparable data to the custodian agencies, these have to provide detailed information regarding inter alia definitions, classifications, methods of computation. The needed information is defined in the IAEG-SDGs Template for Metadata Sustainable Development Goals.

11. The IAEG-SDG can appeal to the custodian agencies to revise their corresponding metadata and give guidance on needed revisions as specified under Article 12. Further, the IAEG can ask international expert groups as for example the Expert Group on International Statistical Classifications or city groups to provide guidance on needed revisions of metadata. Pending the revision and the approval of the new metadata by the IAEG-SDG the corresponding indicator and all time series in question are marked as “content under review” in the global SDG database.

12. Revisions of metadata are inter alia required when:

a) the metadata provided by the custodian agency does not include all the necessary information as defined in the IAEG-SDGs Template for Metadata Sustainable Development Goals, or

b) the metadata provided by the custodian agency does not correspond to the indicator or changes the meaning of the indicator, or

c) the international standard provided in the metadata by the custodian agency is neither widely accepted nor internationally established, or

d) The methods (i.e. method of computation, calculation of aggregates) provided in the metadata by the custodian agency are not internationally established or do not meet statistical standards regarding inter alia soundness and transparency.

13. In case a national data provider finds that the metadata provided by the custodian agency lacks information needed to make national data available or other entities find that the metadata provided by the custodian agency lacks information necessary to understand the data provided in the global database, said entities contact the responsible custodian agency. If no agreement can be found in a consultation between the involved parties the Inter-Agency and Expert Group on Sustainable Development Goal Indicators will be called upon.

14. For the custodian agency to be able to carry out the plausibility check, metadata from national data providers are necessary. In cases where no metadata is made available by the national data provider or the metadata provided does no not include the necessary information as for example stated in the Template for Metadata Sustainable Development Goals, the custodian agency will contact the national data provider with the request to provide the necessary metadata within a reasonable timeframe. If no metadata is provided to the responsible custodian agency by the national data provider in the set timeframe and following two subsequent reminders of the request, the custodian agency may decide to not utilise the corresponding data.

15. Technical assistance may be provided to countries in areas where data are not available, of low quality or differ due to different methodologies and definitions, in order to strengthen national data systems and therefore enable them to produce their own data in the long run.

Annex

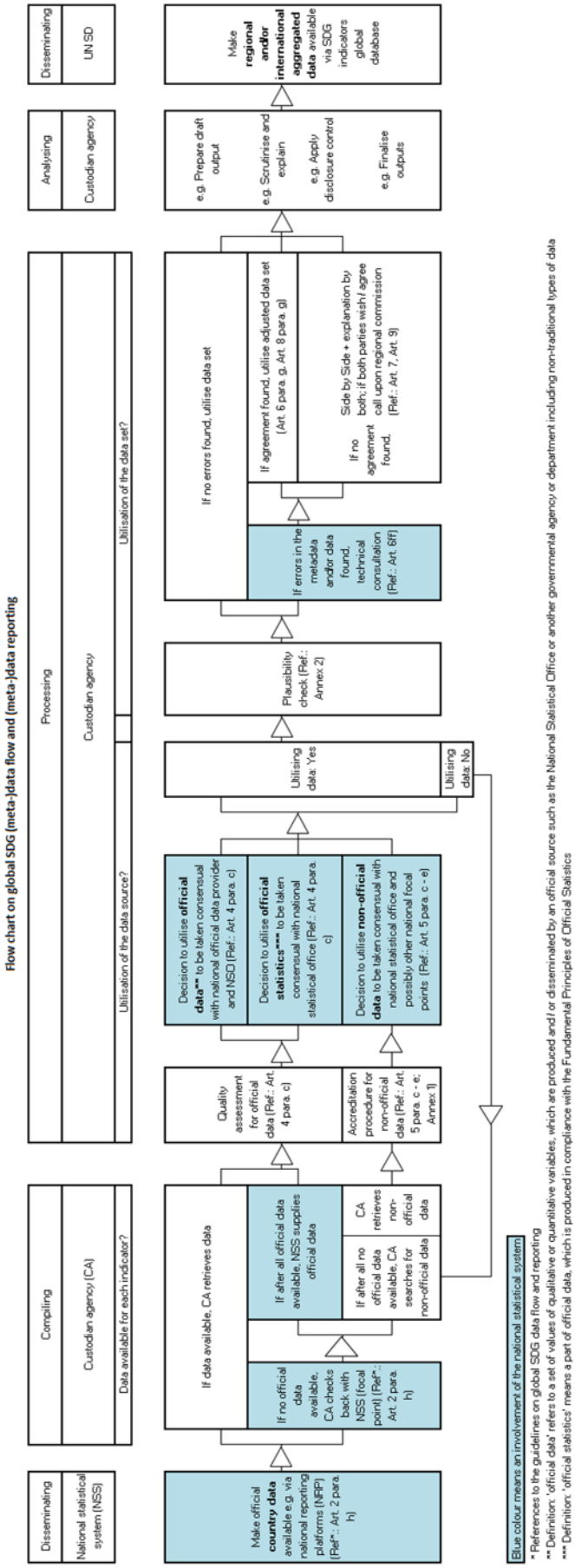
Annex 1: Accreditation Procedure for Non-Official Data (based on "Accreditation procedure for statistical data from non-official sources")

1. Phase 1: Initial examination of sources, data, and metadata
    a) Collection of all freely available information
    b) Circulation of the collected information internally
    c) Examination of the information with a focus on usefulness
2. Phase 2: Acquisition of data and assessment
    a) Closer cooperation with the data provider to acquire essential information such as data set segments, time and method
    b) Update the results of phase 1
    c) Make a first appropriate recommendation, inform the National Statistical Office i.a. including a report on the methodology and provide the possibility to comment
3. Phase 3: Forensic investigation
    a) Producing a clean microdata file
    b) Produce and analyse aggregated statistics
    c) Producing pilot outputs or using the file in the production of existing output
    d) Assessing the capacity of the existing statistical tools to handle the new data
    e) Assess the strengths and weaknesses of the new data
4. Phase 4: Decision
    a) Itemisation of the exact uses of the new data and their impact
    b) Conduct a cost-benefit analysis
    c) Assess risk
    d) Assessment of the feasibility from a legislative and socio-political point of view
    e) Invite the concerned National Statistical Office, the possibly involved national focal point as well as the data set provider to participate in a joint consultation with the aim to find a consensual agreement.
5. Phase 5: Formal Agreement with the data provider (e.g. a Memorandum of Understanding)
    a) Secure cooperation and arrive at a formal and comprehensive agreement
    b) Articulate governance, including change management

Annex 2: Plausibility Checklist (based on "Practical Guide on Data Validation and Editing")

1. Checking the data format.
2. Checking for incomplete data, e.g. missing data.
3. Checking the classification the indicator is based on.
4. Detection of differences between definitions in national and international statistics.
5. Checking for changes in the definitions and classifications used such as concepts, methodologies, surveyed population, data processing over time.
6. Detection of non-admissible values.
7. Detection of incoherencies among variables, i.e. are there relationships and restrictions among variables that have to be satisfied?
8. Detection of large differences between the country's current and past values.
9. Search for breaks in the series, i.e. large jumps or differences in the data from one period to the next.
10. Search for large changes in the series length i.e. the number of observations in a data series supplied by a Member State suffers an important change.
11. Check that aggregated items correspond to the sum of sub-items – when the country provides the breakdown of a given data, the total has to equal the sum of the parts.
12. Cross checking with other sources – the data from a given country should be checked for coherence with other data from the same country or with data from another country.

Annex 3: Data Flow Chart



C. Description of Existing Practices for Global SDG Data Reporting

I. The complementary roles of national and international statistical institutions in the global monitoring

(1) National statistical systems and the coordination role of National Statistical Offices (Excerpt of: Williams, 2016)

The role of National Statistical Systems is to collect data and produce a comprehensive set of integrated statistics, which can be utilised for international aggregates. Therefore national statistics systems will be pivotal in reporting progress towards the SDGs, with the 2030 Agenda clearly recognising the importance of country-led evaluations and data to inform follow-up and review processes at all levels. National Statistical Systems thereby vary in structure. They are usually comprised of national official agencies, especially the National Statistical Office as well as other national data providers.

The primary function of the National Statistical Office is to collect, compile and release official statistics that are produced according to the Fundamental Principles of Official Statistics (FPOS). As stated in the FPOS this includes reliability, objectivity, relevance, statistical confidentiality, transparency, specificity and proportionality. In other words the NSO has to play a crucial role in assuring the quality of official data, in particular official statistics. Depending on several factors, including government priorities and resource availability – both financial and human in terms of availability and capacity, NSOs can take furthermore different roles in assuring the quality of non-official data.

Full Passive Approach – The NSO plays no role at all in the provision of non-official data for the SDGs. Under this approach, the NSO provides no advice, guidelines, or training related to quality for the production of data for indicators that do not use the data from the National Statistical Office. Additionally, under this approach the NSO would not conduct any vetting or review of the data or methodologies being used by non-NSO data producers (other government departments or organisations such as UN organisations, civil society or private sector) for the indicators being used to measure the targets.

The advantage of this approach is that there are no financial costs to the NSO, and the NSO has no responsibility for this data. The disadvantage to this approach is that it may lead to double production of information and inconsistent data or the data used may not be of high quality and may mislead policy makers and result in policies and programs that do not meet the needs of the country. Additionally, under Agenda 2030, countries will be held accountable and that accountability will be linked to data reported. If countries (and NSOs) choose not to provide any validation or quality guidelines for the production and dissemination of data, they cannot be certain about the quality of the information being used

to report against goals, but will nevertheless be held accountable for the indicators reported. For NSOs, despite not producing these data, they may be seen as the data provider and responsible for the data underlying these indicators.

Reactive Approach – NSOs would provide some type of advice, quality guidelines, or training for non-NSO data underlying the SDG indicators only when requested by a data producer. Under this approach, the NSO would only react when requested, but would not actually review data.

The advantage of this approach is that the NSO's role is voluntary and as such, costs can be controlled within existing funding and capacity envelopes. However, it is important to note that there would be sunk costs as quality assurance guidelines and procedures related to non-official data would have to be developed. The disadvantages of this approach are in many respects the same as noted in the Full Passive Approach, because the NSO would examine data only when requested by the data producer. This will likely result in a good proportion of the data not having been validated by the NSO.

Neutral Stance – Under this approach, the NSO would develop quality assurance guidelines, tools and training only. Quality Assurance Guidelines/tools would be made widely available (i.e. online) and training sessions for data producers would take place on an ad-hoc basis. The NSO would not provide or conduct vetting on non-official data.

The advantage of this is that the costs are known upfront and can be easily planned. Additionally, the NSO has made QA tools and information available and would provide training when requested and can be seen as being an advocate for data quality. The NSO itself faces little risk if the quality of non-NSO information is not of good quality. However, as noted under the first option it is likely that the NSO will be considered the data source and/or be called upon to answer questions related to the data. At the country level, the risk is high as the country will be held accountable for the resulting indicator, regardless of the quality of the data.

Proactive Practical Approach – This approach includes wide dissemination and promotion of quality assurance guidelines, tools and training. In addition to this the NSO would provide Quality assurance of *some* non-official produced data. While not acting as an enforcer of data quality, the NSO would examine the metadata, the data and methodology and provide comments and/or recommendations on how to improve data quality. How to determine which non-official data is to be assessed can be done in various ways. For example, data assessment could be based upon a risk assessment (i.e. what is seen as the "riskiest" data could be considered a priority). Alternatively, assessment of non-official data could be conducted based upon national priorities. In this scenario if a country has a focus on a

certain goal or domain, quality assurance assessments could be conducted on non-official data underlying these goals. Another option would be to provide assessment of data coming from non-government sources (such as the private sector, academia, civil society or international organisations). Finally, a mixture of assessment options is possible as each one is not mutually exclusive.

The advantage of this approach is that the NSO is able to assess non-official data and promote quality in a proactive manner. It also enables the NSO to be aware of problem data. Despite not taking on a “policing” role or forcing recommendations to be implemented, under this option the NSO would be in a better position to be able to respond if questioned about the non-NSO data underlying indicators and can note that recommendations to improve data quality were provided to the data producer. Additionally, domestically under this approach, the NSO would be able to advise policy departments about non-official assessed data underlying indicators. This would help ensure effective decision making and good governance. The disadvantages of this approach include costs. Appropriate financial and HR resources will be required in order to ensure that data and costs for assessing data and methodologies are not insignificant – this would require setting up some a unit that can take on these activities. Additionally, should some non-official data be found to be problematic, recommendations for remediation not be followed – it could reflect poorly as the NSO knew there were issues and did not ensure they were remedied. Additionally, at the Government level, the risk remains that it is accountable to answer on progress of SDGs even when the measurement may not be appropriate.

Proactive Custodian – Under this approach the NSO takes on a full custodian role for the underlying data of SDG indicators. It provides, quality assurance guides, tools and training. Moreover, it assesses all non-official data being used for national measurement with respect to SDG indicators. It provides recommendations to improve data quality and ensures that remediation measures are applied and that all data (official, in particular official statistical and non-official data) for the SDG indicators produced by multilateral organisations, other departments, etc. meet minimum data quality standards prior to release.

While under this approach the NSO has full control and can provide reasonable assurance about the non-official data used for SDG measurement, there are significant costs. In some countries it would require a team of trained data quality specialists to adequately vet the data and the timeliness of release of SDG information. Additionally, it will likely be very difficult to restrict non-official data producers from releasing data not deemed of sufficient quality by the NSO.

(2) Responsibilities of data custodian agencies and partner agencies for the global SDG indicators

Custodian Agency

An agency is referred to as “custodian” when the agency either has an existing mandate for global monitoring of the indicator in question and established data reporting mechanisms from countries or when the agency does not have an explicit mandate from its intergovernmental process, but has a well-established global data reporting mechanism, is recognized by Member States, and has a clearly identified counterpart in countries’ government agencies.

The main tasks of custodian agencies are two-fold, first to compile comparable international data series, calculate global and regional aggregates and to provide them, along with the metadata, to the United Nations Statistics Division. A second, equally important task of the custodian agencies is to prepare the storyline for the annual global progress report. In this sense the custodian agencies are working together with the IAEG-SDGs on the implementation of indicators.

Additionally, custodian agencies coordinate the indicator development, in particular for those indicators noted as Tier III, together with National Statistical Systems, other international agencies and stakeholders. This means also to support an increased adoption and compliance with internationally agreed standards.

Against this background it is important that custodian agencies communicate and coordinate with National Statistical Systems in a transparent manner, including on the validation of estimates and data adjustments when these take place.

Finally, custodian agencies also strengthen national statistical capacities through technical assistance and statistical capacity building activities within countries and regions.

Partner Agency

An agency is referred to as “partner agency” when it does not have an explicit mandate for global monitoring, but has a well-established data reporting mechanism for some countries or it has been involved in the methodology development of the indicator.

Responsibilities include:

- Work with data custodian agency on indicator development;
- Facilitate the compilation of the international data series;

- Support increased adoption and compliance with internationally agreed standards; and
- Support countries in strengthening their national statistical capacity.

(3) Role of the Regional Commissions

The Regional Commissions are the regional outposts of the United Nations. Their main objectives are to foster economic integration, to promote the regional implementation of internationally agreed development goals and to support regional sustainable development. In order to achieve these objectives, they promote multilateral dialogue, knowledge sharing and networking at the regional level, and work together to promote intra- regional and inter-regional cooperation. In regards to the Agenda 2030 they also are involved in the monitoring process in their respective regions and to support the global monitoring process. Other tasks in this context are:

- Analysing data and statistical gaps in measuring SDGs;
- Developing (when necessary) regional indicators to complement the global SDG indicators;
- Managing technical cooperation programs for strengthen capacity to provide data on SDG indicators;
- Facilitating peer learning through the sharing of national experiences and best practices for producing statistics for SDGs.

II. Data reporting channels and data flow scenarios from national to global system

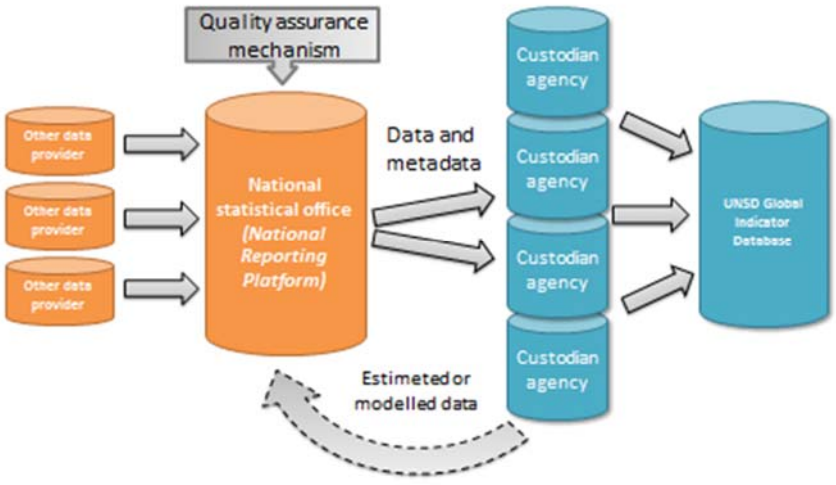
(1) Lessons learnt from the MDG reporting

Since the process of data flows and data reporting for the Sustainable Development Goals is still developing, lessons have to be drawn from experiences of the MDG data flow and reporting. The report of the Task Team on Lessons Learned from MDG Monitoring of the IAEG-MDG, for example, suggests incorporating data quality into the new framework.

(2) Existing data reporting mechanisms (Extract of: UNECE (2017): Guidelines for national SDG-indicators reporting mechanisms, draft)

Depending on the character of the National Statistical System, countries will need to choose a model for how to approach the reporting of the global indicators from the national perspective. National policy plays a fundamental role, especially with respect to NSOs. Circumstances and policies within the countries' National Statistical Systems will determine which model fits best. One country may want to apply a model where all SDG data are channelled through a certain body (usually the NSO) before being sent to the custodian agencies. Such a model may benefit from using an NRP to facilitate collection, quality assurance and transmissions to the receiving agency. When responsibilities for official statistics, including for quality assurance, are dispersed throughout the system, special tools may not be required as some means of "soft coordination" will fit the system better.

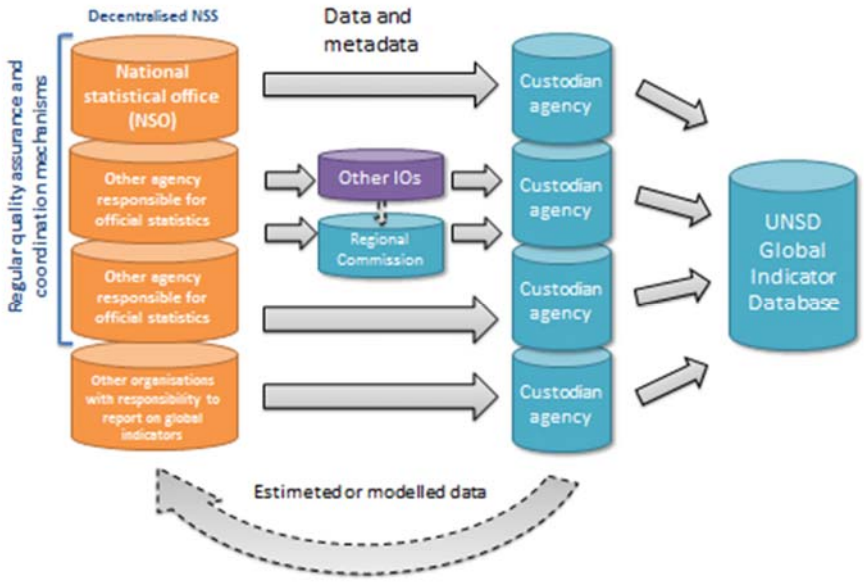
Centralised reporting of statistics for global SDG indicators



In a centralised model of providing SDG indicators, the NSO will typically collect and store all data to a centrally held database. It could also apply a mechanism for validating the data collected (which may include delegations to national statistical programs), and send or otherwise make available the indicators to the custodian agencies and other users. This model could be applied to all indicators, or be limited to statistical indicators only.

Closely associated with this model is the notion of accountability, quality assurance and coordinated accessibility, a one-stop-shop for SDG data. The most important precondition for the model to be efficient is that the NSO has a legally recognised mandate to coordinate and validate all official statistics in the country. The model will also require resources to build and maintain the procedures for collecting, validating and disseminating the global SDG indicators. If this is not already an integral part of the NSOs operations, it could prove costly.

Decentralised reporting of statistics for global SDG indicators



Some countries may not be in favour of centralised collection and reporting of SDG indicators. They would prefer to maintain the responsibility for the production, development and dissemination of official statistics dispersed over many agencies or line ministries. In such cases it could be also more efficient to decentralise the responsibility for providing data on SDG indicators to the entities responsible for producing statistics for the particular indicator. A decentralised model could be associated with the notion of “soft coordination” where the coordination body may issue guidelines and provide trainings and forums for relevant agencies. For this model to be efficient, the coordination relies upon a trusted partnership between statistics producers and on a recognised mandate for the NSO for

certain coordination responsibilities. This model could potentially suffer from lack of overview and accessibility; therefore a set-up like this should consider ways of making SDG data accessible easily and coherently, preferably in an open data format.

“In between” models

The models described above are two extremes. It is highly probable that most countries will choose a model somewhere in between. The quality assurance role of the centralised model could take different forms – such as acting as a “post-office” and simply making data available on NRPs. Other solutions could undertake various degrees of control, from basic validation to full quality control. The role that the NSO takes will typically depend on already existing national policy mechanisms.

In a decentralised model, NSOs can also apply different levels of coordination. The level of ambition may range from simply keeping track of who supplies statistics for which indicators to which custodian agencies, to coordinating all collection, transmission and quality assurance related to the SDG indicators. The role could potentially include, for example:

- serving as a platform/forum for discussion of issues on data collection and analysis between government agencies and international organisations on SDG indicators;
- keeping stakeholders abreast of and share knowledge on statistical activities in the field of data collection and analysis;
- organising and promoting coordination and joint advocacy activities around data collection with a specific focus on SDGs;
- ensuring coordination of information exchange on SDG indicators; and
- promoting substantive discussion on statistical capacity building and coordinating quality assurance activities within the National Statistical System**.**

III. Production of international data series

(1) Review of practices on producing adjusted, estimated and modelled country level data in global reporting

The production of international statistics such as for example in the MDG monitoring process has shown that the comparability and the consistency of reported national data are essential. Differences in the production process can result in the necessity to adjust national statistical data. In order to minimise these adjustments many different organisations and authorities started to define and describe “the standard” statistical process. These standards aim to compare and benchmark processes within and between

organisations and therefore allow better decisions regarding the production process. The following examples provide best practices on producing adjusted, estimated and modelled country level data:

The United Nations Economic Commission for Europe introduced the Generic Statistical Business Process Modell (GSBPM) in 2010. The GSBPM breaks down the data production process into its basic steps. The model further defines and describes each step in the statistical data production process and has become the standard in many statistical organisations around the world. Hopes are that using the same methods and tools during the data production process enables enhanced comparability of data outputs and enhanced data quality. The model encompasses the data production steps: conceptualisation, collection, processing, dissemination and evaluation.

Linked to the GSBPM is the Generic Statistical Data Editing Models (GSDEM), which transposes step five in the GSBPM into a flow model. In the end the input data can be validated, the appropriate method selected and then a possible error corrected.

Furthermore the Practical Guide for Data Validation and Editing describes the standard process of Eurostat on checking the plausibility of received data. For different types of data (micro, country, aggregation) it maps out procedures and methods to identify missing values, invalid or inconsistent entries and potential errors.

Subsequently it describes also the procedures and methods to correct missing values, invalid or inconsistent entries and potential errors according to the different types of data. Furthermore it identifies responsibilities and gives a basic approach on how to deal with these problems amongst National and European Statistical Offices.

References

Eurostat: *Handbook on Data Validation in Eurostat.* https://webgate.ec.europa.eu/fpfis/mwikis/essvalidserv/images/a/ad/PRACTICAL_GUIDE_TO_DATA_VALIDATION.pdf [Retrieved: 08.11.2017]

Eurostat (2014)*: D5: Accreditation procedure for statistical data from non-official sources.*

Expert Group on National Quality Assurance Framework (2012): *Guidelines for the Template for a Generic National Quality Assurance Framework*.

General Assembly of the United Nations (2014): A/RES/68/261 *Fundamental Principles of Official Statistics.*

General Assembly of the United Nations (2015): A/RES/70/1 *Transforming our world: the 2030 Agenda for Sustainable Development.*

General Assembly of the United Nations (2016): A/RES/71/313 *Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable development*.

Independent Expert Advisory Group on a Data Revolution for Sustainable Development (2014) A world that counts.

Inter-agency and Expert Group on Sustainable Development Goal Indicators (2016): *Draft Template for Metadata Sustainable Development Goals.*

United Nations Economic Commission to Europe (2011): *Report on the differences between national and international data for the Millennium Development Goal 7.*

United Nations Economic Commission to Europe (2017): *Draft Guidelines for national SDG-indicators reporting mechanisms.*

United Nations Economic and Social Council (2006): *Strengthening statistical capacity.* 2006.

United Nations Economic and Social Council (2016): E/CN.3/2017/8 *Report of the Secretary-General on quality assurance in the global statistical system.*

United Nations Statistical Commission (UNSC) (2017): *Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development*.

United Nations Statistics Division: *UN Classifications Registry.* https://unstats.un.org/unsd/cr/registry/default.asp?Lg=1 [Retrieved 08.11.2017]

Williams, Cara (2016): *The role of the National Statistical Office in the Agenda 2030 Sustainable Development Goal data and information ecosystem.*