

C.3. Reference metadata

C.3. Reference metadata

18.29. The following items are typically part of the reference metadata associated with statistics on the international supply of services:

- (a) *Legal framework and institutional arrangements*: references to relevant laws and regulations, the role of all institutions involved in compilation and the description of the coordination of the dissemination of statistics and data-sharing agreements among those institutions, either distinctly or as part of broader statistics (e.g., BOP and other external sectors statistics);
- (b) *Underlying concepts and definitions*: definition of residence, non-residence, residence of units, as applicable, definition of statistical value, scope of statistics on the international supply of services and their relationship to national accounts and international merchandise trade statistics, distinction from other international transactions and classification under relevant services item according to BPM6 and EBOPS 2010 and any deviations from international standards, if any, the ultimate controlling institutional unit (UCI) concept, the definition of a foreign affiliate, the definition of direct or indirect control and definitions of statistical and reporting units, etc.;
- (c) *Description of core data sources*: ITRS, enterprise/establishment surveys, surveys of households or persons, administrative records, statistical models, partner country data or a combination of sources, including specific notes on services categories or activities for which particular data collection arrangements or a combination of sources are employed and comments on limitations of source data in terms of coverage, frequency, level of detail, reliability and availability, etc.;
- (d) *Description of data collection, data compilation methods and data-processing procedures*, including the frequency of data collection, the description of specific procedures used for data collection, validation, editing and aggregation, etc., adjustments made to source data, such as imputations, misclassification, adjustments for non-response or under-coverage, adjustments to standard data processing procedures, such as coding, tabulation errors, etc. and indications of departures from international standards, if any;
- (e) *Estimation methods*, such as descriptions of methods for estimating non-reported transactions or transactions falling below customs and/or ITRS thresholds (e.g., cost, insurance and freight (CIF)-free on board (FOB) adjustments for the transportation item);
- (f) *Dissemination policy*, including release and revision schedules, an indication of the presentation format of data, the level of disaggregation and eventual commentaries accompanying the data, etc.;
- (g) *Additional explanations and footnotes concerning the data as required*: explanatory notes on revisions, breaks in series, definitions of confidentiality flags, etc.;
- (h) *Quality reporting*, including the publication of regular quality reports that use the quality dimensions in the Template for a Generic National Quality Assurance Framework (NQAF)^[1] and include definitions of such quality dimensions as timeliness, accessibility and comparability;
- (i) *Confidentiality*, including descriptions of confidentiality rules and indications of how much data is affected by such rules.^[2]

18.30. **Compilation of metadata** Metadata are compiled at all stages of the statistics production process. The present *Guide* encourages countries to use the following good practices, as applicable, in metadata compilation:

- (a) *Use standardized metadata concepts*. In the same way as any data item, metadata items must also be clearly defined. Even though each statistical domain, including statistics of international trade in services, has its specific metadata items, it is good practice to use applicable standardized concepts that are relevant across statistical domains (e.g., by adopting cross-domain concepts from the SDMX framework or the OECD Glossary of Statistical Terms). The aim should be to promote the harmonization of statistical information and their related high-level metadata across various institutions and statistical domains, even if some specific metadata concepts are not applicable or are organized differently in different domains or institutions;
- (b) *When developing metadata for statistics compiled within the framework for describing the international supply of services, use the metadata developed in the related statistical domains and used in your country*. Statistics of international trade in services is a relatively new statistical domain in many countries. However, it is very likely that the metadata policy is already in place in related statistical domains. Compilers are advised to carefully review and use such metadata;
- (c) *Define layers of metadata*. It is good practice to compile metadata in layers of incremental detail and provide clear links between high-level and specific metadata concepts. Such a layered structure of metadata will allow data users and analysts to access necessary metadata items and to minimize the risk of misinterpretation of data content when, for example, compiling data from various data sources. It will also ensure the clear presentation of metadata to diverse groups of users;
- (d) *Establish metadata registries*. A metadata registry is a central repository (usually a database itself) with information that allows for the linking of the detailed definitions (semantics) with the codes (representations) of the metadata items used to describe a particular statistical data set. It is good practice for compilers to put special emphasis on the development, maintenance and dissemination of metadata registries to improve the harmonization, standardization, use, reuse and exchange of their metadata;^[3]
- (e) *Confidentiality and access to metadata during the compilation process*. As metadata for statistics on the international supply of services might be compiled by various units of the same agency or by units located in different organizations, there might be cases in which metadata describes data that is subject to confidentiality rules. It is good practice, in that context, for confidentiality rules to be set up in a such way that they will allow compilers to obtain non-confidential data aggregates with the same metadata content;
- (f) *Incorporate structural metadata items into the data processing as early as possible (e.g., as parts of the records structure)*. That step will facilitate data processing, including the identification of viable options for data aggregation and subsequent presentation. It is advisable for structural metadata to be made an integral part of the database containing statistics compiled within the framework for describing the international supply of services in such a way that it can be extracted together with any data item and used in data processing to obtain meaningful combined data sets;

(g) *Establish clear links between data and metadata.* As metadata are generated and processed during every step of the data compilation process, there is a strong requirement to ensure that the appropriate metadata retain their links with data. In that connection, it is good practice to implement metadata-driven management in the various stages of the statistical production process;^[4]

(h) *Compilation of reference metadata.* Reference metadata can be presented as a detailed explanatory note describing the scope, coverage, and quality of a data set and can be made available electronically alongside the database or in special publications.

18.31. **Priorities in metadata management** Although, ideally, the management of metadata would take into account all the recommendations highlighted thus far, countries with less developed systems for statistics on the international supply of services should begin by setting up an exhaustive, consistent and detailed repository (possibly in the form of a *metadata registry*) with both structural and reference metadata, adopting as much as possible standardized cross-domain metadata concepts, both nationally and internationally. The next immediate priority should be granting to the general public easy, extensive and timely access to metadata. In subsequent phases, the system could be improved by the gradual incorporation of more advanced features, such as a layered presentation of metadata and active links between data and metadata.

Back to C. Indicative lists of metadata items

[1] See <https://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx>; and see chapter 19 for more information on the national quality assurance framework (NQAF) and quality reporting and management in general.

[2] See chapter 20 for more information on statistical confidentiality.

[3] The Euro-SDMX registry includes harmonized structural metadata, the DSDs designed for the statistical domains and metadata structure definitions, e. g., Euro-SDMX Metadata Structure (ESMS) and other related information.

[4] There are several information model specifications that can contribute to achieving that goal (most notably SDMX and the Data Documentation Initiative (DDI), which are designed to perform different functions, but can be used together in the same system, or complement each other in the compilation and exchange of data and metadata.