

A. Basic concepts (Chapter 25)

25.2. *Definition and role of metadata.* Metadata are “data about data” which enable and facilitate sharing, querying, understanding and using statistical data over the different stages of collection, compilation and dissemination, and at their various levels of aggregation (i.e., from microdata to macrodata). They encompass administrative facts about the data (who has created them and when), and definition of concepts applied, as well as description of how data were collected and processed before they were disseminated or stored in a database (see IMTS, para. 9.20). Not only are metadata important for users, but they also play a crucial role in the statistical production process, as common standards and definitions should be followed to the extent possible throughout all statistical domains in order to facilitate the linking and integration of statistical information.

25.3. *Structural metadata.* Structural metadata are identifiers and descriptors which are essential for discovering, organizing, retrieving and processing statistical data sets.^[1] They can be thought of as the “labels” that need to be associated to each data item in order for it to have any meaning at all. In the context of international merchandise trade statistics, structural metadata include items such as unit of measurement, time period, commodity code, identification of reporting and partner countries (country codes), identification of trade flow, etc. Structural metadata also include the information required to link data points across periods in order to obtain meaningful time series.

25.4. *Reference metadata.* Reference metadata are of a more general nature and “may refer to specific statistical data, to entire data collections or even to the institution that provides the data”.^[2] More specifically, in the context of international merchandise trade statistics, reference metadata items include: (a) explanations of the concepts and definitions adopted and their practical implementation (e.g., coverage and valuation); (b) details on the methodologies used for the generation of the data (e.g., specification of data source(s), description of the sampling framework in the context of survey-based data, description of data validation and editing techniques, etc.); and (c) information describing the various quality dimensions of the resulting international merchandise trade statistics (e.g., relevance, accuracy, timeliness, methodological soundness, coherence and accessibility).^[3]

25.5. *Metadata and data quality.* There is a bidirectional relationship between metadata and data quality. On the one hand, metadata provide details on the various quality dimensions of international merchandise trade statistics, as stated in the previous paragraph. On the other hand, the availability of adequate metadata to users is in itself a quality component through its role as an indicator of the accessibility of international merchandise trade statistics (see IMTS 2010, paras. 9.17 and 9.21). Compilers should aim to provide users with all the metadata required to understand both the strengths and the limitations of the international merchandise trade statistics they produce, documenting in a timely manner all methodological aspects underlying the data that are relevant for their proper use and interpretation (e.g., definitions, classifications, scope, confidentiality issues, etc.).

25.6. *Institutional arrangements for metadata compilation.* To reduce the burden associated with trade metadata projects, it is good practice for compilers to closely cooperate with the specific units responsible for ensuring within the national statistical system that metadata are produced, that they adhere to a standard format and that they are properly maintained and updated.

[1] See Statistical Data and Metadata Exchange, “SDMX content-oriented guidelines” (2009), annex 4, entitled “Metadata common vocabulary”.

[2] Ibid., p.6

[3] Ibid., annex 4.