

## 2. The role of a statistical metadata system



Unknown macro: 'scrollbar'

### What is a statistical metadata system?

Metadata can be defined as "data that define and describe other data" whereas statistical metadata are "data about statistical data, and comprise data and other documentation that describe objects in a formalised way" (both definitions come from the 2009 edition of the SDMX Metadata Common Vocabulary). Some metadata, however, describe other resources that are relevant to the collection, processing and dissemination of data (for example, questionnaires and publications) rather than relating directly to data themselves, so a slightly wider definition of metadata may need to be considered. With this in mind, the SDMX Metadata Common Vocabulary definition of a Statistical Metadata System (SMS) can be used: "A data processing system that uses, stores and produces statistical metadata". The term system refers to the people, processes and technology involved in managing statistical metadata.

Metadata have two basic functions. The first is to uniquely and formally define the content and links between objects and processes in the statistical information system. The second is to determine all related technical parameters.

With rapidly developing technologies, it can be challenging to implement efficient strategies for the production and dissemination of statistics. The growing use of the Internet has caused significant changes in the priorities for managing statistical metadata. In the past, priority was often given to technical metadata and information technology challenges, whereas now there has been a clear shift to focusing on content and methodological issues. When designing the SMS, priority should be given to the content of the system. The primary focus should be achieving business outcomes and ensuring the expected benefits can be realised through effective governance and processes.

Due to these changes, transparent and integrated descriptions of information flows within and outside the statistical office are vital. The use of technology for data collection, interactive communication with users, and dissemination of statistics, calls for a coherent and well functioning metadata system.

The implementation of an SMS should be independent of the technology used for statistical data processing. However, the links between the SMS and electronic processing systems must be ensured. Processing of statistical data should be driven by metadata stored in the SMS.

### What is the role of the statistical metadata system?

The success of an SMS can be measured by the extent to which the needs of diverse groups of statistical metadata users are satisfied. The need for metadata is defined by the various activities, tasks and processes carried out in a statistical organization. Therefore, the role of SMS should be understood in the context of these processes and activities.

A basic framework for the role of the SMS in statistical organizations is defined in:

- [The Fundamental Principles of Official Statistics](#)  
(adopted by the United Nations Statistical Commission in April 1994); and
- [The Principles Governing International Statistical Activities](#)  
(endorsed by the Committee for Coordination of Statistical Activities in September 2005).

The principles relevant to metadata management are highlighted below:

#### The Fundamental Principles of Official Statistics

**Principle 2.** To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.

**Principle 3.** To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.

**Principle 5.** Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.

**Principle 8.** Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.

**Principle 9.** The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.

## The Principles Governing International Statistical Activities

**Principle 1.** High quality international statistics, accessible for all, are a fundamental element of global information systems.

**Principle 2.** To maintain the trust in international statistics, their production is to be impartial and strictly based on the highest professional standards.

**Principle 4.** Concepts, definitions, classifications, sources, methods and procedures employed in the production of international statistics are chosen to meet professional scientific standards and are made transparent for the users.

**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.

**Principle 7.** Erroneous interpretation and misuse of statistics are to be immediately appropriately addressed.

[Back to top](#)

The SMS should be a tool enabling a statistical organization to effectively perform the following functions:

1. Planning, designing, implementing and evaluating statistical production processes.
2. Managing, unifying and standardizing workflows and processes.
3. Documenting data collection, storage, evaluation and dissemination.
4. Managing methodological activities, standardizing and documenting concept definitions and classifications.
5. Managing communication with end-users of statistical outputs and gathering of user feedback.
6. Improving the quality of statistical data and transparency of methodologies. Ensuring and evaluating the quality of statistical data is one of the most important activities. To this end, national and international statistical organizations have adopted a set of criteria (relevance and completeness, comparability and coherence of statistical concepts, accuracy of statistical estimations, timeliness and punctuality of delivered statistical information, its accessibility and clarity). The SMS should offer a relevant set of metadata for all of these criteria.
7. Managing statistical data sources and cooperation with respondents.
8. Improving discovery and exchange of data between the statistical organization and its users.
9. Improving integration of statistical information systems with other national information systems. Growing demands to use administrative data for statistical purposes require better integration and sharing of metadata between statistical and administrative bodies, to ensure coherence and consistency of exchanged information.
10. Disseminating statistical information to end users. End users need reliable metadata for searching, navigation, and interpretation of data. Metadata should also be available to assist postprocessing of statistical data.
11. Improving integration between national and international organizations. International organizations are increasingly requiring integration of their own metadata with metadata of national statistical organizations in order to make statistical information more comparable and compatible, and to monitor the use of agreed standards.
12. Developing a knowledge base on the processes of statistical information systems, to share knowledge among staff and to minimize the risks related to knowledge loss when staff leave or change functions.
13. Improving administration of statistical information systems, including administration of responsibilities, compliance with legislation, performance and user satisfaction.
14. Facilitating the evaluation of costs and revenues for the statistical organization.
15. Unifying statistical terminology as a vehicle for better communication and understanding between managers, designers, subject-matter statisticians, methodologists, respondents and users of statistical information systems.

[Back to top](#)