

RECOMMENDATIONS

COMMISSION

COMMISSION RECOMMENDATION

of 23 June 2009

on reference metadata for the European Statistical System

(Text with EEA relevance)

(2009/498/EC)

THE COMMISSION OF THE EUROPEAN COMMUNITIES,

Having regard to the Treaty establishing the European Community, and in particular Article 211 thereof,

Whereas:

- (1) The European Statistics Code of Practice ⁽¹⁾, intended for the national and Community statistical authorities enumerates 15 principles linked to the institutional environment, statistical processes and statistical output.
- (2) Principle 15 of the European Statistics Code of Practice deals with accessibility and clarity of European Statistics, also emphasising that the accompanying metadata should be documented according to a standardised metadata system.
- (3) Reference metadata are an integral part of the metadata system of each statistical authority.
- (4) With the adoption of the European Statistics Code of Practice the national and Community statistical authorities have committed themselves towards high quality statistics which also requires more transparent and harmonised reporting on data quality.

- (5) In the framework of the SDMX initiative on common technical and statistical standards for the exchange and sharing of data and metadata launched by the Bank of International Settlements, the European Central Bank, the Community statistical authority (Eurostat), the International Monetary Fund, the Organisation for Economic Cooperation and Development, the United Nations and the World Bank the SDMX Content-Oriented Guidelines have been established which support the creation and implementation of harmonised reference metadata in the European Statistical System.

- (6) Considerable efficiency gains can be reached when the reference metadata are produced on the basis of a harmonised list of statistical concepts within the European Statistical System, while at the same time allowing national and Community statistical authorities to add more statistical concepts in particular statistical areas if needed.
- (7) Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics ⁽²⁾ provides a reference framework for this Recommendation,

HEREBY RECOMMENDS TO THE MEMBER STATES:

1. National statistical authorities are invited to apply the statistical concepts and sub-concepts listed in the Annex when reference metadata are compiled in the different statistical areas and when reference metadata are exchanged within the European Statistical System or beyond.

⁽¹⁾ Commission Recommendation of 25 May 2005 on the independence, integrity and accountability of the national and Community statistical authorities, COM(2005) 217 final.

⁽²⁾ OJ L 87, 31.3.2009, p. 164.

2. Additional statistical concepts should be added by national statistical authorities to the above list of concepts and sub-concepts if it is necessary for particular statistical areas.
3. National statistical authorities are invited to regularly inform the Commission (Eurostat) of their application of the concepts and sub-concepts listed in the Annex.

Done at Brussels, 23 June 2009.

For the Commission
Joaquín ALMUNIA
Member of the Commission

ANNEX

List of statistical concepts and sub-concepts (including the definitions of the concepts and sub-concepts)

| Number | Concepts | Sub-concepts | Descriptions |
|--------|--------------------------|--------------------------------------|---|
| 1. | Contact | | Individual or organisational contact points for the data or metadata, including information on how to reach the contact points. |
| 1.1. | | Contact organisation | The name of the organisation of the contact points for the data or metadata. |
| 1.2. | | Contact organisation unit | An addressable subdivision of an organisation. |
| 1.3. | | Contact name | The name of the contact points for the data or metadata. |
| 1.4. | | Contact person function | The area of technical responsibility of the contact, such as 'methodology', 'database management' or 'dissemination'. |
| 1.5. | | Contact mail address | The postal address of the contact points for the data or metadata. |
| 1.6. | | Contact e-mail address | E-mail address of the contact points for the data or metadata. |
| 1.7. | | Contact phone number | The telephone number of the contact points for the data or metadata. |
| 1.8. | | Contact fax number | Fax number of the contact points for the data or metadata. |
| 2. | Metadata update | | The date on which the metadata element was inserted or modified in the database. |
| 2.1. | | Metadata last certified | Date of the latest certification provided by the domain manager to confirm that the metadata posted are still up to date, even if the content has not been amended. |
| 2.2. | | Metadata last posted | Date of the latest dissemination of the metadata. |
| 2.3. | | Metadata last update | Date of last update of the content of the metadata. |
| 3. | Statistical presentation | | |
| 3.1. | | Data description | Main characteristics of the data set described in an easily understandable manner, referring to the data and indicators disseminated. |
| 3.2. | | Classification system | Arrangement or division of objects into groups based on characteristics which the objects have in common. |
| 3.3. | | Sector coverage | Main economic or other sectors covered by the statistics. |
| 3.4. | | Statistical concepts and definitions | Statistical characteristics of statistical observations. |
| 3.5. | | Statistical unit | Entity for which information is sought and for which statistics are ultimately compiled. |
| 3.6. | | Statistical population | The total membership or population or 'universe' of a defined class of people, objects or events. |

| Number | Concepts | Sub-concepts | Descriptions |
|--------|----------------------------|----------------------------------|---|
| 3.7. | | Reference area | The country or geographic area to which the measured statistical phenomenon relates. |
| 3.8. | | Time coverage | The length of time for which data are available. |
| 3.9. | | Base period | The period of time used as the base of an index number, or to which a constant series refers. |
| 4. | Unit of measure | | The unit in which the data values are measured. |
| 5. | Reference period | | The period of time or point in time to which the measured observation is intended to refer. |
| 6. | Institutional Mandate | | Set of rules or other formal set of instructions assigning responsibility as well as the authority to an organisation for the collection, processing, and dissemination of statistics. |
| 6.1. | | Legal acts and other agreements | Legal acts or other formal or informal agreements that assign responsibility as well as the authority to an agency for the collection, processing, and dissemination of statistics. |
| 6.2. | | Data sharing | Arrangements or procedures for data sharing and coordination between data producing agencies. |
| 7. | Confidentiality | | A property of data indicating the extent to which their unauthorised disclosure could be prejudicial or harmful to the interest of the source or other relevant parties. |
| 7.1. | | Confidentiality — policy | Legislative measures or other formal procedures which prevent unauthorised disclosure of data that identify a person or economic entity either directly or indirectly. |
| 7.2. | | Confidentiality — data treatment | Rules applied for treating the data set to ensure statistical confidentiality and prevent unauthorised disclosure. |
| 8. | Release policy | | Rules for disseminating statistical data to interested parties. |
| 8.1. | | Release calendar | The schedule of statistical release dates. |
| 8.2. | | Release calendar access | Access to the release calendar information. |
| 8.3. | | User access | The policy for release of the data to users, the scope of dissemination (e.g. to the public, to selected users), how users are informed that the data are being released, and whether the policy determines the dissemination of statistical data to all users. |
| 9. | Frequency of dissemination | | The time interval at which the statistics are disseminated over a given time period. |
| 10. | Dissemination format | | Media by which statistical data and metadata are disseminated. |
| 10.1. | | News release | Regular or ad-hoc press releases linked to the data. |
| 10.2. | | Publications | Regular or ad-hoc publications in which the data are made available to the public. |
| 10.3. | | Online database | Information about online databases in which the disseminated data can be accessed. |

| Number | Concepts | Sub-concepts | Descriptions |
|--------|--------------------------------|------------------------------|---|
| 10.4. | | Micro-data access | Information on whether micro-data are also disseminated. |
| 10.5. | | Other | References to the most important other data dissemination done. |
| 11. | Accessibility of documentation | | |
| 11.1. | | Documentation on methodology | Descriptive text and references to methodological documents available. |
| 11.2. | | Quality documentation | Documentation on procedures applied for quality management and quality assessment. |
| 12. | Quality Management | | Systems and frameworks in place within an organisation to manage the quality of statistical products and processes. |
| 12.1. | | Quality assurance | All systematic activities implemented that can be demonstrated to provide confidence that the processes will fulfil the requirements for the statistical output. |
| 12.2. | | Quality assessment | Overall assessment of data quality, based on standard quality criteria. |
| 13. | Relevance | | The degree to which statistical information meet current and potential needs of the users. |
| 13.1. | | User needs | Description of users and their respective needs with respect to the statistical data. |
| 13.2. | | User satisfaction | Measures to determine user satisfaction. |
| 13.3. | | Completeness | The extent to which all statistics that are needed are available. |
| 14. | Accuracy and reliability | | Accuracy: closeness of computations or estimates to the exact or true values that the statistics were intended to measure. Reliability: closeness of the initial estimated value to the subsequent estimated value. |
| 14.1. | | Overall accuracy | Assessment of accuracy, linked to a certain data set or domain, which is summarising the various components. |
| 14.2. | | Sampling error | That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a subset of the population is enumerated. |
| 14.3. | | Non-sampling error | Error in survey estimates which cannot be attributed to sampling fluctuations. |
| 15. | Timeliness and punctuality | | |
| 15.1. | | Timeliness | Length of time between data availability and the event or phenomenon they describe. |
| 15.2. | | Punctuality | Time lag between the actual delivery of the data and the target date when it should have been delivered. |

| Number | Concepts | Sub-concepts | Descriptions |
|--------|------------------------|------------------------------|--|
| 16. | Comparability | | Measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas or over time. |
| 16.1. | | Comparability — geographical | The extent to which statistics are comparable between geographical areas. |
| 16.2. | | Comparability over time | The extent to which statistics are comparable or reconcilable over time. |
| 17. | Coherence | | Adequacy of statistics to be reliably combined in different ways and for various uses. |
| 17.1. | | Coherence — cross domain | The extent to which statistics are reconcilable with those obtained through other data sources or statistical domains. |
| 17.2. | | Coherence — internal | The extent to which statistics are consistent within a given data set. |
| 18. | Cost and burden | | Cost associated with the collection and production of a statistical product and burden on respondents. |
| 19. | Data revision | | Any change in a value of a statistic released to the public. |
| 19.1. | | Data revision — policy | Policy aimed at ensuring the transparency of disseminated data, whereby preliminary data are compiled that are later revised. |
| 19.2. | | Data revision — practice | Information on the data revision practice. |
| 20. | Statistical processing | | |
| 20.1. | | Source data | Characteristics and components of the raw statistical data used for compiling statistical aggregates. |
| 20.2. | | Frequency of data collection | Frequency with which the source data are collected. |
| 20.3. | | Data collection | Systematic process of gathering data for official statistics. |
| 20.4. | | Data validation | Process of monitoring the results of data compilation and ensuring the quality of the statistical results. |
| 20.5. | | Data compilation | Operations performed on data to derive new information according to a given set of rules. |
| 20.6. | | Adjustment | The set of procedures employed to modify statistical data to enable it to conform to national or international standards or to address data quality differences when compiling specific data sets. |
| 21. | Comment | | Supplementary descriptive text which can be attached to data or metadata. |