SLOVENSKI STANDARD
SIST ISO 19114:2003

01-december-2003

Geografske informacije - Postopki za ocenjevanje kakovosti

Geographic information -- Quality evaluation procedures

Information géographique -- Procédures d'évaluation de la qualité

Ta slovenski standard je istoveten z: **ISO 19114:2003**

ICS:

03.120.99	Drugi standardi v zvezi s kakovostjo	Other standards related to quality
07.040	Astronomija. Geodezija. Geografija	Astronomy. Geodesy. Geography
35.240.70	Uporabniške rešitve IT v znanosti	IT applications in science

SIST ISO 19114:2003 **en**

INTERNATIONAL STANDARD

ISO 19114

First edition
2003-08-15

Geographic information — Quality evaluation procedures

Information géographique — Procédures d'évaluation de la qualité

Reference number
ISO 19114:2003(E)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 19114 was prepared by Technical Committee ISO/TC 211, *Geographic information/Geomatics*.

Introduction

For the purpose of evaluating the quality of a dataset, clearly defined procedures must be used in a consistent manner. This enables data producers to express how well their product meets the criteria set forth in its product specification and enables data users to establish the extent to which a dataset meets their requirements. The quality of a dataset is described using two components: a quantitative component and a non-quantitative component. The objective of this International Standard is to provide guidelines for evaluation procedures of quantitative quality information for geographic data in accordance with the quality principles described in ISO 19113. It also offers guidance on reporting quality information.

This International Standard recognizes that a data producer and a data user may view data quality from different perspectives. Conformance quality levels can be set using the data producer's product specification or a data user's data quality requirements. If the data user requires more data quality information than that provided by the data producer, the data user may follow the data producer's data quality evaluation process flow to get the additional information. In this case, the data user requirements are treated as a product specification for the purpose of using the data producer process flow.

The quality evaluation procedures described in this International Standard, when applied in accordance with ISO 19113, provide a consistent and standard manner to determine and report the quality information in a dataset.

Geographic information — Quality evaluation procedures

1 Scope

This International Standard provides a framework of procedures for determining and evaluating quality that is applicable to digital geographic datasets, consistent with the data quality principles defined in ISO 19113. It also establishes a framework for evaluating and reporting data quality results, either as part of data quality metadata only, or also as a quality evaluation report.

This International Standard is applicable to data producers when providing quality information on how well a dataset conforms to the product specification, and to data users attempting to determine whether or not the dataset contains data of sufficient quality to be fit for use in their particular applications.

Although this International Standard is applicable to all types of digital geographic data, its principles can be extended to many other forms of geographic data such as maps, charts and textual documents.

2 Conformance

This International Standard defines three classes of conformance: one for quality evaluation procedures, one for evaluating data quality, and one for reporting quality information. The abstract test suites for the three classes of conformance are given in Annex A.

3 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 19113:2002, *Geographic information — Quality principles*

ISO 19115:2003, *Geographic information — Metadata*

4 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 19113 and ISO 19115 (some of which are repeated for convenience) and the following apply.

4.1
conformance quality level
threshold value or set of threshold values for data quality results used to determine how well a dataset meets the criteria set forth in its product specification or user requirements

4.2
dataset
identifiable collection of data

[ISO 19115]

NOTE A dataset may be a smaller grouping of data which, though limited by some constraint such as spatial extent or feature type, is located physically within a larger dataset. For purposes of data quality evaluation, a dataset may be as small as a single feature or feature attribute contained within a larger dataset.

4.3
dataset series
collection of datasets sharing the same product specification

[ISO 19115]

4.4
direct evaluation method
method of evaluating the quality of a dataset based on inspection of the items within the dataset

4.5
full inspection
inspection of every item in a dataset

NOTE Full inspection is also known as 100 % inspection.

4.6
indirect evaluation method
method of evaluating the quality of a dataset based on external knowledge

NOTE Examples of external knowledge are dataset lineage, such as production method or source data.

4.7
item
that which can be individually described or considered

[ISO 2859-1]

NOTE An item can be any part of a dataset, such as a feature, feature relationship, feature attribute, or combination of these.

4.8
population
totality of items under consideration

[ISO 3534-2]

EXAMPLE 1 All points in a dataset.

EXAMPLE 2 Names of all roads in a certain geographic area.

4.9
reference data
data accepted as representing the universe of discourse, to be used as reference for direct external quality evaluation methods

5 Abbreviated terms

ADQR aggregated data quality results

AQL acceptance quality limit [ISO 3534-2]

RMSE root mean square error

6 Process for evaluating data quality

6.1 General

A quality evaluation process may be used in different phases of a product life cycle, having different objectives in each phase. The phases of the life cycle considered here are specification, production, delivery, use and update. Annex B describes some specific dataset-related operations to which quality evaluation procedures are applicable.

The process for evaluating data quality is a sequence of steps to produce and report a data quality result. A quality evaluation process consists of the application of quality evaluation procedures to specific dataset-related operations performed by the dataset producer and the dataset user.

Processes for evaluating data quality are applicable to static datasets and to dynamic datasets. Dynamic datasets are datasets that receive updates so frequently that for all practical purposes they are continuously changing. Annex C describes the application of the process to evaluate data quality to dynamic datasets.

6.2 Components of the process

6.2.1 Process flow

The quality evaluation process is a sequence of steps taken to produce a quality evaluation result. Figure 1 illustrates the process flow for evaluating and reporting data quality results.

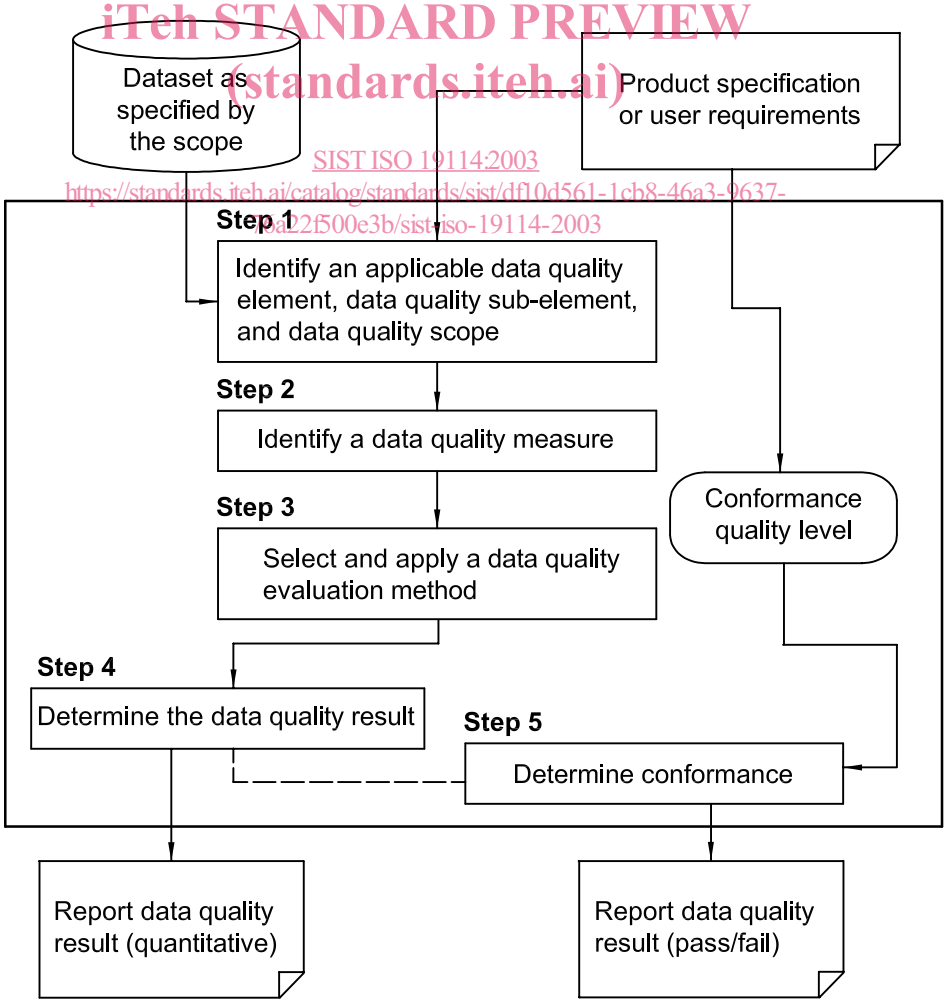

Figure 1 — Evaluating and reporting data quality results

6.2.2 Process steps

Table 1 specifies the process steps.

Table 1 — Process steps

Process step	Action	Description
1	Identify an applicable data quality element, data quality sub-element, and data quality scope	The data quality element, data quality sub-element, and data quality scope to be tested is identified in accordance with the requirements of ISO 19113. This is repeated for as many different tests as required by the product specification or user requirements.
2	Identify a data quality measure	A data quality measure, data quality value type and, if applicable, a data quality value unit is identified for each test to be performed. Annex D provides examples of data quality measures for the data quality elements and data quality sub-elements given in ISO 19113. Annex D, by these examples, provides assistance to the user in selection of a measure.
3	Select and apply a data quality evaluation method	A data quality evaluation method for each identified data quality measure is selected. NOTE A spatial description of the results (achievable by spatial interpolation of the results, graphical portrayal, etc.) might be useful, corresponding not to a result, but to a different, although related, dataset.
4	Determine the data quality result	A quantitative data quality result, a data quality value or set of data quality values, a data quality value unit and a date are the output of applying the method.
5	Determine conformance	Whenever a conformance quality level has been specified in the product specification or user requirements, the data quality result is compared with it to determine conformance. A conformance data quality result (pass-fail) is the comparison of the quantitative data quality result with a conformance quality level.

7 Data quality evaluation methods

7.1 Classification of data quality evaluation methods

A data quality evaluation procedure is accomplished through the application of one or more data quality evaluation methods. Data quality evaluation methods are divided into two main classes: direct and indirect. Direct methods determine data quality through the comparison of the data with internal and/or external reference information. Indirect methods infer or estimate data quality using information on the data, such as lineage. The direct evaluation methods are further subclassified by the source of the information needed to perform the evaluation. Figure 2 depicts this classification structure.

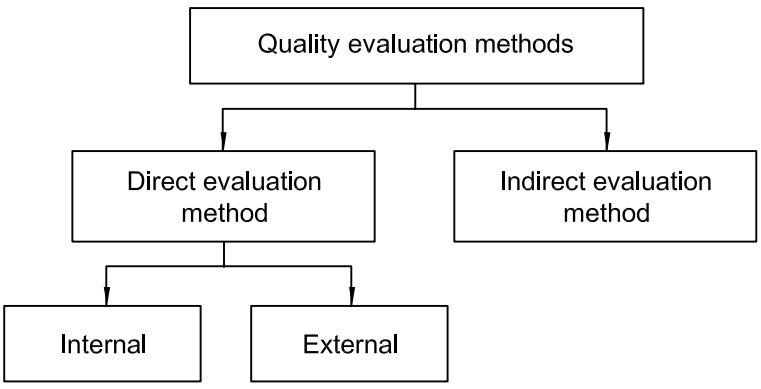

Figure 2 — Classification of data quality evaluation methods (informative)

7.2 Direct evaluation methods

7.2.1 Types of direct evaluation methods

The direct evaluation method is further subdivided into internal and external. All the data needed to perform an internal direct data quality evaluation method are internal to the dataset being evaluated.

EXAMPLE 1 All the data necessary to perform a logical consistency test for topological consistency of boundary closure resides in a topologically structured dataset.

External direct quality evaluation requires reference data external to the dataset being tested.

EXAMPLE 2 The data needed to perform a completeness test for the road names in a dataset requires another information source of road names.

EXAMPLE 3 A positional accuracy test requires a reference dataset or a new survey.

7.2.2 Means of accomplishing direct evaluation

For both external and internal evaluation methods, there are two considerations, automated or non-automated and full inspection or sampling.

Data quality elements and data quality sub-elements which are easily checked by automated means include the following:

a) logical consistency: format consistency;

   EXAMPLE Check data fields for positive entry.

b) logical consistency: topological consistency;

   EXAMPLE Polygon closure.

c) logical consistency: domain consistency;

   EXAMPLE Bounds violations, specified domain value violations.

d) completeness: omission;

   EXAMPLE Comparison check of street names from another file.

e) completeness: commission;

   EXAMPLE Comparison check of street names from another file.

f) temporal accuracy: temporal consistency.

   EXAMPLE Check all records for appropriate range of dates.

7.2.3 Full inspection

Full inspection requires testing every item in the population specified by the data quality scope. Table 2 describes the procedure for full inspection that shall be used.

Table 2 — Procedure for full inspection

Procedure step	Description
Define items	An item is a minimum unit to be inspected. An item can be a feature, a feature attribute or a feature relationship.
Inspect items in the data quality scope	Inspect every item it the data quality scope.

NOTE Full inspection is most appropriate for small populations or for tests that can be accomplished by automated means.

7.2.4 Sampling

Sampling requires testing sufficient items in the population in order to achieve a data quality result. Table 3 describes the sampling procedure that shall be used.

Table 3 — Sampling procedure

Procedure step	Description
Define a sampling method	Examples of sampling methods are given in Annex E. Those methods include simple random sampling, stratified sampling (e.g. guided by feature type, a feature relationship or an area), multistage sampling and non-random sampling.
Define items	An item is a minimum unit to be inspected. An item can be a feature, a feature attribute or a feature relationship.
Divide data quality scope (population) into lots	A lot is a collection of items in the data quality scope from which a sample is drawn and inspected. Each lot should, as far as possible, consist of items produced under the same conditions and at the same time.
Divide lots to sampling units	Sampling unit is the area of the lot where inspection is conducted.
Define the sampling ratio or sample size	A sampling ratio gives information on how many items on average are extracted for inspection from each lot.
Select sampling units	Select required number of sampling units so that the sampling ratio or sample size for items is fulfilled.
Inspect items in the sampling units	Inspect every item in the sampling units.

The sampling procedure shall be reported in accordance with Clause 8.

The ISO 2859 series and ISO 3951-1 may be applied to sampling for evaluating conformance to a product specification. These standards were originally developed for non-spatial use. Annex E of this International Standard gives examples describing how to apply the ISO 2859 series and ISO 3951-1 and provides guidelines on how to define samples and devise sampling methods, taking the geographic nature of the data into account.

The reliability of the data quality result should be analysed when using sampling, especially when using small sample sizes and methods other than simple random sampling.

7.3 Indirect evaluation method

The indirect evaluation method is a method of evaluating the quality of a dataset based on external knowledge. This external knowledge may include, but is not limited to, data quality overview elements and other quality reports on the dataset or data used to produce the dataset.

NOTE 1 This method is recommended only if direct evaluation methods cannot be used.

NOTE 2 Usage information records uses of a dataset. This is helpful when searching for datasets that have been produced or used for specific purposes.

NOTE 3 Lineage information records information about the production and history of the dataset. It includes information about, for example, source materials to produce a dataset or the production processes applied. This is useful when determining the suitability of a dataset for a given use. An example is lineage metadata relating to a digital terrain model file that has been created by means of stereo-correlation from images captured under certain conditions. Experience tells the evaluator that the horizontal positional RMSE is 10 m for this type of imagery. Or, for example, lineage metadata of a digitized 1:25 000 scale topographic map indicates conformance to a town planner's requirements for a base map.

NOTE 4 Purpose information describes the purpose for which the dataset was produced. A purpose may be in support of a specific requirement, or the dataset may be a general purpose dataset for several uses. This is useful when identifying the possible value of a dataset.

7.4 Data quality evaluation examples

Examples of typical methods used and how they may be applied are described in Annexes F, G and H.

8 Reporting data quality evaluation information

8.1 Reporting as metadata

Quantitative quality results shall be reported as metadata in compliance with ISO 19115, which contains the related model and data dictionary.

8.2 Reporting in a quality evaluation report

There are two conditions under which a quality evaluation report shall be produced:

a) when data quality results reported as metadata are only reported as pass/fail;

b) when aggregated data quality results are generated.

The report is required in the latter condition to explain how aggregation was done and how to interpret the meaning of the aggregate result. However, a quality evaluation report may be created at any other time (such as to provide more detail than reported as metadata) but a quality evaluation report cannot be used in lieu of reporting as metadata.

A quality evaluation report shall be produced in compliance with Annex I which contains the relevant model and data dictionary.

8.3 Reporting aggregated data quality result

When several quality results are aggregated into a single quality result for reporting the quality of a dataset, the aggregated data quality result shall be reported as metadata and shall be included in the data quality report. The data quality result shall be reported as type “aggregate”. Annex J describes the production of aggregate data quality results and Annex H provides a production example.