भारतीय मानक
Indian Standard

IS/ISO/IEC 11179-4 : 2004

[Superseding IS 16638 (Part 4) : 2018]

सूचना प्रोद्यौगिकी — मेटाडेटा रजिस्ट्रीज (एमडीआर)

भाग 4 डेटा परिभाषा का निर्धारण

Information Technology — Metadata Registries (MDR)

Part 4 Formulation of Data Definitions

ICS 35.040.50

भारतीय मानक ब्यूरो
BUREAU OF INDIAN STANDARDS
मानक भवन, 9 बहादुर शाह ज़फर मार्ग, नई दिल्ली - 110002
MANAK BHAVAN, 9 BAHADUR SHAH ZAFAR MARG
NEW DELHI - 110002
www.bis.gov.in www.standardsbis.in

December 2022 **Price Group 7**

NATIONAL FOREWORD

This Indian Standard (Part 4) which is identical with ISO/IEC 11179-4 : 2004 'Information technology — Metadata registries (MDR) — Part 4: Formulation of data definitions' issued by International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) jointly, was adopted by the Bureau of Indian Standards on the recommendations of the Data Management System Sectional Committee, LITD 15, and approval of the Electronics and Information Technology Division Council.

This standard was originally published in 2008 as IS 16638 (Part 4) : 2018 and was identical to ISO/IEC 11179-4 : 2004. The supersession of this Standard has been undertaken to align it with the latest version of ISO/IEC 11179-4 : 2004 under single numbering. On publication of this Standard IS 16638 (Part 4) : 2018 Stand withdrawn.

The other parts of ISO/IEC 11179 series are already published as Indian standards as IS/ISO/IEC 11179-x series.

The text of ISO/IEC Standard has been approved as suitable for publication as an Indian Standard without deviations. Certain conventions are, however, not identical to those used in Indian Standards. Attention is particularly drawn to the following:

a) Wherever the words 'International Standard' appear referring to this standard, they should be read as 'Indian Standard'.

b) Comma (,) has been used as a decimal marker while in Indian Standards, the current practice is to use a point (.) as the decimal marker.

For the purpose of deciding whether a particular requirement of this standard is complied with, the final value, observed or calculated, expressing the result of a test or analysis, shall be rounded off in accordance with IS 2 : 2022 'Rules for rounding off numerical values (*second revision*)'. The number of significant places retained in the rounded off value should be the same as that of the specified value in this standard.

Contents

Page

Introduction

Data processing and electronic data interchange rely heavily on accurate, reliable, controllable and verifiable data recorded in databases. A prerequisite for correct and proper use and interpretation of data is that both users and owners of data have a common understanding of the meaning and representation of the data. To facilitate this common understanding, a number of characteristics, or attributes of the data have to be defined. These characteristics of data are known as “metadata”, that is, “data that describes data”. This part of ISO/IEC 11179 specifies requirements and recommendations on the formulation of data definitions that are specified in Metadata Registries. The purpose of these definitions is to specify, describe, explain, and clarify the meaning of data, to promote the standardization or reuse of data elements, and to promote data sharing and integration of information systems.

The structure of a *Metadata Registry* is specified in the form of a conceptual data model. The *Metadata Registry* is used to keep information about data elements and associated concepts, such as “data element concepts”, “conceptual domains”, and “value domains”. Generically, these are all referred to as “metadata items”. Such metadata are necessary to clearly describe, record, analyse, classify, and administer data.

The definitional requirements and recommendations specified in this part of ISO/IEC 11179 do not always apply to terminological definitions found in glossaries and language dictionaries. Differences exist between the requirements that apply in a language dictionary, and the requirements that apply in a metadata registry. The requirements for ISO/IEC 11179 are more restrictive than those for a natural language dictionary. For example, a language dictionary may have multiple definitions covering multiple senses of a term or word, whereas data definitions are developed for particular contexts and should not have multiple senses within any context. Data definitions are intended to explicate the concept or concepts, which are represented by a collection of data, a data value, a data element, or other metadata item. A single definition may be established as the reference definition, with other definitions asserted to be equivalent (e.g., a definition in one language may be established as a reference definition, with definitions in other languages asserted to be equivalent). Metadata items may have a single preferred definition within a particular context, with other deprecated definitions.

Many data definitions include terms that themselves need to be defined (e.g., “charge”, “allowance”, “delivery”). Some of these terms may have different definitions in different industrial sectors. Therefore, there is a need for most metadata registries to establish an associated *glossary or terminology reference* of terms used in the definitions.

Indian Standard

INFORMATION TECHNOLOGY — METADATA REGISTRIES (MDR)
PART 4 FORMULATION OF DATA DEFINITIONS

1 Scope

This part of ISO/IEC 11179 specifies requirements and recommendations for constructing definitions for data and metadata. Only semantic aspects of definitions are addressed; specifications for formatting the definitions are deemed unnecessary for the purposes of ISO/IEC 11179. While especially applicable to the content of metadata registries as specified in ISO/IEC 11179-3, this part of ISO/IEC 11179 is useful broadly for developing definitions for data and metadata.

These definitional requirements and recommendations pertain to formulating definitions for data elements and other types of data constructs such as entity types, entities, relationships, attributes, object types (or classes), objects, composites, code entries, metadata items, and the data referred to by XML tags.

2 Conformance

This part of ISO/IEC 11179 may be used independently, e.g., for defining data elements outside the context of a metadata registry. In such cases, compliance may be claimed if the requirements and recommendations have been followed in developing the definitions.

Where used in the context of an ISO/IEC 11179 metadata registry, this part of ISO/IEC 11179 shall constitute the criteria for definitions when establishing the registration status as specified in ISO/IEC 11179-6.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1
Conceptual Domain
set of valid value meanings

NOTE The value meanings may either be enumerated or expressed via a description.

[ISO/IEC 11179-3:2003, 3.3.21]

3.2
Concept
unit of knowledge created by a unique combination of characteristics

[ISO 1087-1:2000, 3.2.1]

3.3
data
re-interpretable representation of information in a formalized manner suitable for communication, interpretation or processing

NOTE Data can be processed by human or automatic means.

[ISO/IEC 2382-1:1993, 01.01.02]

3.4
data element
unit of **data** for which the definition, identification, representation and permissible values are specified by means of a set of attributes

[ISO/IEC 11179-3:2003, 3.3.36]

3.5
Data Element Concept
concept that can be represented in the form of a data element, described independently of any particular representation

[ISO/IEC 11179-3:2003, 3.3.38]

3.6
definition
representation of a concept by a descriptive statement which serves to differentiate it from related concepts

[ISO 1087-1:2000, 3.3.1]

3.7
metadata
data that defines and describes other data

[ISO 1087-1:2000, 3.2.18]

3.8
Metadata Registry
information system for registering metadata

[ISO/IEC 11179-3:2003, 3.2.22]

3.9
metadata item
instance of a **metadata object**

NOTE 1 In all parts of ISO/IEC 11179, this term is applied only to instances of metadata objects described by the metamodel in Clause 4 of ISO/IEC 11179-3:2003. Examples include instances of data elements, data element concepts, permissible values, etc.

NOTE 2 A metadata item has associated attributes, as appropriate for the metadata object it instantiates.

[ISO/IEC 11179-3:2003, 3.2.19]

3.10
metadata object
object type defined by a metamodel

NOTE In all parts of ISO/IEC 11179, this term is applied only to metadata objects described by the metamodel in Clause 4 of ISO/IEC 11179-3:2003. Examples include data elements, data element concepts, permissible values etc. See Clause 3.3 of ISO/IEC 11179-3:2003 for a complete list.

[ISO/IEC 11179-3:2003, 3.2.20]

3.11
name
designation of an object by a linguistic expression

[ISO/IEC 11179-3:2003, 3.2.26]

3.12
Value Domain
set of permissible values

NOTE 1 The value domain provides representation, but has no implication as to what data element concept the values may be associated with nor what the values mean.

NOTE 2 The permissible values may either be enumerated or expressed via a description.

[ISO/IEC 11179-3:2003, 3.3.140]

4 Summary of data definition requirements and recommendations

A listing of the requirements and recommendations without explanations is provided in this clause for convenience of the user. The intent is to facilitate ease of use of this document once an understanding of the requirements and recommendations is achieved. Clause 5 describes each requirement and recommendation with an explanation and examples to ensure their exact meaning is understood.

4.1 Requirements

A data definition shall:

a) be stated in the singular

b) state what the concept is, not only what it is not

c) be stated as a descriptive phrase or sentence(s)

d) contain only commonly understood abbreviations

e) be expressed without embedding definitions of other data or underlying concepts

4.2 Recommendations

A data definition should:

a) state the essential meaning of the concept

b) be precise and unambiguous

c) be concise

d) be able to stand alone

e) be expressed without embedding rationale, functional usage, or procedural information

f) avoid circular reasoning

g) use the same terminology and consistent logical structure for related definitions

h) be appropriate for the type of metadata item being defined

5 Provisions

5.1 Premises

Data is used for specific purposes. Differences in use require different operational manifestations of some requirements and recommendations. For example, different levels of specificity for data definitions are generally required in different contexts. Recommendation 5.3.a) below, provides an example of this need for varying levels of specificity for different definitions. The implementation of Recommendation 5.3.a), "state the essential meaning of the concept" is context dependent. The primary characteristics deemed necessary to convey the essential meaning of a particular definition will vary according to the level of generalization or specialization of the data. Primary and essential characteristics for defining concepts such as "airport" in the commercial air transportation industry might be specific, where a more general definition may be adequate in a different context. Within a metadata registry, multiple equivalent definitions may be written in different languages or, within a single language, for different audiences such as children, general public, or subject area specialists. For a discussion of relationships between concepts in different contexts and how characteristics are used to differentiate concepts, see ISO 704, Clause 5. Definitions should be written to facilitate understanding by any user and by recipients of shared data.

5.2 Requirements

To facilitate understanding of the requirements for construction of well-formed data definitions, explanations and examples are provided below. Each requirement is followed by a short explanation of its meaning. Examples are given to support the explanations. In all cases, a good example is provided to exemplify the explanation. When deemed beneficial, a poor, but commonly used example is given to show how a definition should **NOT** be constructed. To further explain the differences between the good and poor examples, examples are followed by a statement of rationale behind them. Note that the examples below are definitions for data elements and these definitions are illustrative.

A data definition shall:

a) **be stated in the singular**

EXPLANATION - The concept expressed by the data definition shall be expressed in the singular. (An exception is made if the concept itself is plural.)

EXAMPLE - "Article Number"

1) good definition: A reference number that identifies an article.

2) poor definition: Reference number identifying articles.

REASON - The poor definition uses the plural word "articles," which is ambiguous, since it could imply that an "article number" refers to more than one article.

b) **state what the concept is, not only what it is not**

EXPLANATION - When constructing definitions, the concept cannot be defined exclusively by stating what the concept is **not.**

EXAMPLE - "Freight Cost Amount"

1) good definition: Cost amount incurred by a shipper in moving goods from one place to another.

2) poor definition: Costs which are not related to packing, documentation, loading, unloading, and insurance.

REASON - The poor definition does not specify what is included in the meaning of the data.

c) **be stated as a descriptive phrase or sentence(s)** (in most languages)

EXPLANATION - A phrase is necessary (in most languages) to form a precise definition that includes the essential characteristics of the concept. Simply stating one or more synonym(s) is insufficient. Simply restating the words of the name in a different order is insufficient. If more than a descriptive phrase is needed, use complete, grammatically correct sentences.

EXAMPLE - “Agent Name”

1) good definition: Name of party authorized to act on behalf of another party.

2) poor definition: Representative.

REASON - “Representative” is a near-synonym of the data element name, which is not adequate for a definition.

d) **contain only commonly understood abbreviations**

EXPLANATION - Understanding the meaning of an abbreviation, including acronyms and initialisms, is usually confined to a certain environment. In other environments the same abbreviation can cause misinterpretation or confusion. Therefore, to avoid ambiguity, full words, not abbreviations, shall be used in the definition.

Exceptions to this requirement may be made if an abbreviation is commonly understood such as “i.e.” and “e.g.” or if an abbreviation is more readily understood than the full form of a complex term and has been adopted as a term in its own right such as “radar” standing for “radio detecting and ranging.”

All acronyms must be expanded on the first occurrence.

EXAMPLE 1 - “Tide Height”

1) good definition: The vertical distance from mean sea level (MSL) to a specific tide level.

2) poor definition: The vertical distance from MSL to a specific tide level.

REASON - The poor definition is unclear because the acronym, MSL, is not commonly understood and some users may need to refer to other sources to determine what it represents. Without the full word, finding the term in a glossary may be difficult or impossible.

EXAMPLE 2 - “Unit of Density Measurement”

1) good definition: The unit employed in measuring the concentration of matter in terms of mass per unit (m.p.u.) volume (e.g., pound per cubic foot; kilogram per cubic meter).

2) poor definition: The unit employed in measuring the concentration of matter in terms of m.p.u. volume (e.g., pound per cubic foot; kilogram per cubic meter).

REASON - m.p.u. is not a common abbreviation, and its meaning may not be understood by some users. The abbreviation should be expanded to full words.

e) **be expressed without embedding definitions of other data or underlying concepts**

EXPLANATION - As shown in the following example, the definition of a second data element or related concept should not appear in the definition proper of the primary data element. Definitions of terms should be provided in an associated glossary. If the second definition is necessary, it may be attached by a note at the end of the primary definition's main text or as a separate entry in the dictionary. Related definitions can be accessed through relational attributes (e.g., cross-reference).

EXAMPLE 1- “Sample Type Code”

1) good definition: A code identifying the kind of sample.

2) poor definition: A code identifying the kind of sample collected. A sample is a small specimen taken for testing. It can be either an actual sample for testing, or a quality control surrogate sample. A quality control sample is a surrogate sample taken to verify results of actual samples.

REASON - The poor definition contains two extraneous definitions embedded in it. They are definitions of “sample” and of “quality control sample.”

EXAMPLE 2 - "Issuing Bank Documentary Credit Number"

1) good definition: Reference number assigned by issuing bank to a documentary credit.

2) poor definition: Reference number assigned by issuing bank to a documentary credit. A documentary credit is a document in which a bank states that it has issued a documentary credit under which the beneficiary is to obtain payment, acceptance, or negotiation on compliance with certain terms and conditions and against presentation of stipulated documents and such drafts as may be specified.

REASON - The poor definition contains a concept definition, which should be included in a glossary.

5.3 Recommendations

A data definition should:

a) **state the essential meaning of the concept**

EXPLANATION - All primary characteristics of the concept represented should appear in the definition at the relevant level of specificity for the context. The inclusion of non-essential characteristics should be avoided. The level of detail necessary is dependent upon the needs of the system user and environment.

EXAMPLE 1 - “Consignment Loading Sequence Number” (Intended context: any form of transportation)

1) good definition: A number indicating the sequence in which consignments are loaded in a means of transport or piece of transport equipment.

2) poor definition: A number indicating the sequence in which consignments are loaded in a truck.

REASON - In the intended context, consignments can be transported by various transportation modes, e.g., trucks, vessels or freight trains. Consignments are not limited to trucks for transport.

EXAMPLE 2 - “Invoice Amount”

1) good definition: Total sum charged on an invoice.

2) poor definition: The total sum of all chargeable items mentioned on an invoice, taking into account deductions on one hand, such as allowances and discounts, and additions on the other hand, such as charges for insurance, transport, handling, etc.

REASON - The poor definition includes extraneous material.

b) **be precise and unambiguous**

EXPLANATION -The exact meaning and interpretation of the defined concept should be apparent from the definition. A definition should be clear enough to allow only one possible interpretation.

EXAMPLE - “Shipment Receipt Date”

1) good definition: Date on which a shipment is received by the receiving party.

2) poor definition: Date on which a specific shipment is delivered.

REASON - The poor definition does not specify what determines a “delivery.” “Delivery” could be understood as either the act of unloading a product at the intended destination or the point at which the intended customer actually obtains the product. It is possible that the intended customer never receives the product that has been unloaded at his site or the customer may receive the product days after it was unloaded at the site.

c) **be concise**

EXPLANATION - The definition should be brief and comprehensive. Extraneous qualifying phrases such as “for the purpose of this metadata registry,” “terms to be described,” shall be avoided.

EXAMPLE - “Character Set Name”

1) good definition: The name given to the set of phonetic or ideographic symbols in which data is encoded.

2) poor definition: The name given to the set of phonetic or ideographic symbols in which data is encoded, for the purpose of this metadata registry, or, as used elsewhere, the capability of systems hardware and software to process data encoded in one or more scripts.

REASON - In the poor definition, all the phrases after “...data is encoded” are extraneous qualifying phrases.

d) **be able to stand alone**

EXPLANATION - The meaning of the concept should be apparent from the definition. Additional explanations or references should not be necessary for understanding the meaning of the definition.

EXAMPLE - “School Location City Name”

1) good definition: Name of the city where a school is situated.

2) poor definition: See “school site”.

REASON - The poor definition does not stand alone, it requires the aid of a second definition (school site) to understand the meaning of the first.

e) **be expressed without embedding rationale, functional usage, domain information, or procedural information**

EXPLANATION - Although they are often necessary, such statements do not belong in the definition proper because they contain information extraneous to the definition. If deemed useful, such expressions may be placed in other metadata attributes (see ISO/IEC 11179-3). It is, however, permissible to add examples after the definition.

1) The rationale for a given definition should not be included as part of the definition (e.g. if a data element uses miles instead of kilometers, the reason should not be indicated in the definition).

2) Functional usage such as: “this data element should not be used for ...” should not be included in the definition proper.

3) Remarks about procedural aspects. For example, “This data element is used in conjunction with data element 'xxx'”, should not appear in the definition; instead use “Related data reference” and “Type of relationship” as specified in ISO/IEC 11179-3.

EXAMPLE - “Data Field Label”

1) good definition: Identification of a field in an index, thesaurus, query, database, etc.

2) poor definition: Identification of a field in an index, thesaurus, query, database, etc., which is provided for units of information such as abstracts, columns within tables.

REASON - The poor definition contains remarks about functional usage. This information starting with “which is provided for...” must be excluded from the definition and placed in another attribute, if it is necessary information.

f) **avoid circular reasoning**

EXPLANATION - Two definitions shall not be defined in terms of each other. A definition should not use another concept's definition as its definition. This results in a situation where a concept is defined with the aid of another concept that is, in turn, defined with the aid of the given concept.

EXAMPLE - two data elements with poor definitions:

1) Employee ID Number - Number assigned to an employee.

2) Employee - Person corresponding to the employee ID number.

REASON - Each definition refers to the other for its meaning. The meaning is not given in either definition.

g) **use the same terminology and consistent logical structure for related definitions**

EXPLANATION - A common terminology and syntax should be used for similar or associated definitions.

EXAMPLE - The following example illustrates this idea. Both definitions pertain to related concepts and therefore have the same logical structure and similar terminology.

1) “Goods Dispatch Date” - Date on which goods were dispatched by a given party.

2) “Goods Receipt Date” - Date on which goods were received by a given party.

REASON - Using the same terminology and syntax facilitates understanding. Otherwise, users wonder whether some difference is implied by use of synonymous terms and variable syntax.

h) **be appropriate for the type of metadata item being defined**

EXPLANATION – Different types of metadata item in a metadata registry (e.g. data element concept, data element, conceptual domain, value domain) each play a different role and this should be reflected in the definitions.

EXAMPLE –

Data element concept: “Job Grade Maximum Salary Amount”

Definition: The maximum salary permitted for the associated job grade.

Note: The data element concept makes no reference to a specific value domain.

Conceptual Domain: “Monetary amount”

Definition: An amount that may be expressed in a unit of currency.

Note: The definition refers to a “dimensionality” of currency, but not to a specific currency.

Data element 1": “European Job Grade Maximum Salary Amount”

Definition: The maximum salary permitted for the associated job grade expressed in Euros.

Data element 2": “U.S. Job Grade Maximum Salary Amount”

Definition: The maximum salary permitted for the associated job grade expressed in US dollars.

Note: Data element definitions may refer to explicit values domains, since this may be all that distinguishes two data elements.

Bibliography

[1] ISO 704:2000, *Terminology work — Principles and methods*

[2] ISO 1087-1:2000, *Terminology work — Vocabulary — Part 1: Theory and application*

[3] ISO/IEC 2382-1:1993, *Information technology — Vocabulary — Part 1: Fundamental terms*

[4] ISO/IEC 11179-3:2003, *Information technology — Metadata registries (MDR) — Part 3: Registry metamodel and basic attributes*

Bureau of Indian Standards

BIS is a statutory institution established under the *Bureau of Indian Standards Act*, 2016 to promote harmonious development of the activities of standardization, marking and quality certification of goods and attending to connected matters in the country.

Copyright

Review of Indian Standards

Amendments are issued to standards as the need arises on the basis of comments. Standards are also reviewed periodically; a standard along with amendments is reaffirmed when such review indicates that no changes are needed; if the review indicates that changes are needed, it is taken up for revision. Users of Indian Standards should ascertain that they are in possession of the latest amendments or edition by referring to the website-www.bis.gov.in or www.standardsbis.in.

This Indian Standard has been developed from Doc No.: LITD 15 (18170).

Amendments Issued Since Publication

Amend No.	Date of Issue	Text Affected

BUREAU OF INDIAN STANDARDS

Headquarters:

Manak Bhavan, 9 Bahadur Shah Zafar Marg, New Delhi 110002
Telephones: 2323 0131, 2323 3375, 2323 9402 *Website*: www.bis.gov.in

Regional Offices: *Telephones*

Central : 601/A, Konnectus Tower -1, 6th Floor, DMRC Building, Bhavbhuti Marg, New Delhi 110002 { 2323 7617

Eastern : 8th Floor, Plot No 7/7 & 7/8, CP Block, Sector V, Salt Lake, Kolkata, West Bengal 700091 { 2367 0012 2320 9474

Northern : Plot No. 4-A, Sector 27-B, Madhya Marg, Chandigarh 160019 { 265 9930

Southern : C.I.T. Campus, IV Cross Road, Taramani, Chennai 600113 { 2254 1442 2254 1216

Western : Plot No. E-9, Road No.-8, MIDC, Andheri (East), Mumbai 400093 { 2821 8093

Branches : AHMEDABAD. BENGALURU. BHOPAL. BHUBANESHWAR. CHANDIGARH. CHENNAI. COIMBATORE. DEHRADUN. DELHI. FARIDABAD. GHAZIABAD. GUWAHATI. HIMACHAL PRADESH. HUBLI. HYDERABAD. JAIPUR. JAMMU & KASHMIR. JAMSHEDPUR. KOCHI. KOLKATA. LUCKNOW. MADURAI. MUMBAI. NAGPUR. NOIDA. PANIPAT. PATNA. PUNE. RAIPUR. RAJKOT. SURAT. VISAKHAPATNAM.

Published by BIS, New Delhi