TECHNICAL SPECIFICATION

ISO/TS 21526

First edition
2019-10

Health informatics — Metadata repository requirements (MetaRep)

Informatique de santé — Exigences relatives aux référentiels de métadonnées (MetaRep)

Reference number
ISO/TS 21526:2019(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: 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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 215, *Health informatics*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

This document is intended to be an extension to and a clarification of the ISO/IEC 11179 series.

Healthcare has a fundamental requirement for describing the meaning, provenance and governance of data, and for setting standards for how that data is stored and communicated. Unsurprisingly, it uses metadata registries and repositories extensively for a wide range of purposes supporting care delivery, reporting and research. However, these registries are only partially interoperable, with consequences to cost of ownership and utility that lead to under-use, particularly where simple, read-only repositories are required. There is also considerable unmet need resulting from the community's focus on message-based interoperability at the expense of the description of the meaning of the data in the systems that are the source of that data, and how that data maps to the meaning of the large quantities of data communicated between providers and national bodies in tables or simple XML.

Data in healthcare systems should persist in content and meaning across organizations and time for a wide variety of uses including patient care, patient safety, service management, service improvement, public health and healthcare research. The sharing and adoption of record or message designs offer immediate and tangible benefits to these ends, entailing organizations to adopt common standards for the exchange of the specifications of records and messages in terms of the representation and definition of individual elements of data; the inter-relationships of those elements in data models and where sets of data that accord to those models can be found together with any contextual information about those data sets that is required for their understanding and appropriate reuse.

Settings where metadata collections are assembled include individual clinics and hospitals, organizations managing a portfolio of clinical studies, organizations providing cloud applications in support of healthcare and regional and national bodies who commission standard reports and datasets in pursuit of policy objectives. The intent is to support an ecosystem of interoperable registries and repositories which facilitate both the development and implementation of content standards – and thus interoperable content – and the publication of interoperable metadata about the kinds of data available in both care and research.

This document includes a review of existing components of ISO/IEC 11179-3, i.e. Metadata Registry (MDRMetamodel), and ISO/IEC 19763-12 Metamodel for Information Models (MFIInformationModel) to specify where variations from or additions to the requirements of these standards are required for specific healthcare use cases. Registries conforming to this document are also likely to reference of ISO/IEC 11179-7[1)] Metamodel for dataset registration (MDRDatasets) and ISO/IEC TS 19763-13 Metamodel for forms registration (MFIForms), however it is less clear that simplifications or extensions to either are necessary in the healthcare or healthcare research setting and thus the original documents should be used as is.

MDRMetamodel provides a comprehensive model for an international metadata registry addressing several large communities with overlapping concerns, and the conformance statements in the 2013 edition are framed in this context. Equally MFIInformationModel is designed to represent models represented in many ways from purely conceptual entity relationship diagrams through to a concrete relational database instance. A metadata registry/repository aimed at a less diverse community such as healthcare or directed at the needs of a smaller organisation might not require the complete implementation of ISO/IEC 11179-3 and ISO/IEC 19763-12, so it is important that any restriction or simplification is shared to preserve registry interoperability.

1) Under preparation. Stage at the time of publication: ISO/IEC/PRF 11179-7:2019.

Health informatics — Metadata repository requirements (MetaRep)

1 Scope

This document describes requirements for collections of metadata about data elements and their containing models and datasets in a healthcare environment. The collection can serve as a repository or as dictionary describing a set of items in use in a particular domain, organisation or product for reference, or it can additionally serve as a registry, supporting the development and communication of standard items to an audience with shared goals.

This document specifies standard refinements that account for the detailed governance and administration requirements that are particular to healthcare data. It focuses on the description of data that is persisted in healthcare systems rather than the specification of messages between these systems. It describes necessary extensions to the ISO/IEC 11179 series and to other International Standards on metadata originating from ISO/IEC JTC 1/SC 32 to address the wider variety of value domain types found in healthcare. Where appropriate, it also suggest restrictions/simplifications to the ISO/IEC 11179 series that promote wider adoption without compromising interoperability between metadata registries and opportunities for the development of tooling that consumes metadata for the generation or the parameterization of code.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

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

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
acceptability rating
scale of acceptability

3.2
administered item
registered item for which *administrative information* (3.3) is recorded

3.3
administrative information
<metadata registry> information about the administration of an item in a *metadata registry* (3.33)

EXAMPLE Creation date, last change date, origin, change description, explanatory comment.

3.4
administration record
collection of *administrative information* (3.3) for an *administered item* (3.2)

3.5
administrative status
designation of the status in the administrative process of a *registration authority* (3.42) for handling *registration* (3.41) requests

Note 1 to entry: The values and associated meanings of "administrative status" are determined by each registration authority.

3.6
attribute
characteristic (3.12) of an object or entity

3.7
assertion
sentence or proposition in logic which is asserted (or assumed) to be true

3.8
attribute instance
specific instance of an *attribute* (3.6)

Note 1 to entry: Adapted from ISO 2382-17:1993 to distinguish an instance of an attribute from its value.

[SOURCE: ISO/IEC 11179-3:2013, 3.2.7]

3.9
binding
mapping from one framework or specification to another, enabling *data* (3.20) and/or commands to be passed between them

3.10
boolean
mathematical *datatype* (3.23) associated with two-valued logic

[SOURCE: ISO/IEC 11179-3:2013, 3.2.12]

3.11
cardinality
number of elements in a set

3.12
characteristic
abstraction of a *property* (3.39) of an object or of a set of objects

Note 1 to entry: Characteristics are used for describing concepts.

[SOURCE: ISO/IEC 11179-3:2013, 3.2.14]

3.13
class
description of a set of objects that share the same *attributes* (3.6), operations, methods, *relationships* (3.44), and semantics

3.14
classification scheme
descriptive information for an arrangement or division of objects into groups based on *characteristics* (3.12), which the objects have in common

3.15
concept
unit of knowledge created by a unique combination of *characteristics* (3.12)

[SOURCE: ISO/IEC 11179-3:2013, 3.2.18]

3.16
concept system
set of *concepts* (3.15) structured according to the relations among them

[SOURCE: ISO/IEC 11179-3:2013, 3.2.19]

3.17
conceptual domain
CD
concept (3.15) that expresses its description or valid instance meanings

Note 1 to entry: The value meanings may either be enumerated or expressed via a description.

[SOURCE: ISO/IEC 11179-3:2013, 3.2.21]

3.18
contact
instance of a role of an individual or *organization* (3.37) (or organization part or organization person) to or from whom an information item(s), a material object(s) and/or person(s) can be sent in a specified *context* (3.19)

3.19
context
circumstance, purpose, and perspective under which an object is defined or used

3.20
data
re-interpretable representation of facts, *concepts* (3.15), or instructions in a formalized manner suitable for communication, interpretation, or processing

Note 1 to entry: Data can be processed by human or automatic means.

[SOURCE: ISO/IEC 11179-3:2013, 3.2.27]

3.21
data element
DE
unit of *data* (3.20) for which the *definition* (3.28), identification, representation and *permissible values* (3.38) are specified by means of a set of *attributes* (3.6)

3.22
data element concept
DEC
concept (3.15) that can be represented in the form of a *data element* (3.21), described independently of any particular representation

3.23
datatype
set of distinct *values* (3.46), characterized by properties of those values and by operations on those values

[SOURCE: ISO/IEC 11179-3:2013, 3.1.9]

3.24
dimensionality
expression of measurement without units

Note 1 to entry: A quantity is a value with an associated unit of measure. 32º Fahrenheit, 0º Celsius, $100 USD, and 10 reams (of paper) are quantities. Equivalence between two units of measure is determined by the existence of a quantity preserving one-to-one correspondence between values measured in one unit of measure and values measured in the other unit of measure, independent of context, and where characterizing operations are the same. Equivalent units of measure in this sense have the same dimensionality. The equivalence defined here forms an equivalence relation on the set of all units of measure. Each equivalence class corresponds to a dimensionality. The units of measure "temperature" in degrees Fahrenheit" and "temperature in degrees Celsius" have the same dimensionality, because for each value measured in degrees Fahrenheit there is a value measured in degrees Celsius with the same quantity, and vice-versa. The same operations may be performed on quantities in each unit of measure. Quantity preserving one-to-one correspondences are the well-known equations Cº = (5/9)*(Fº – 32) and Fº = (9/5)*(Cº) + 32.

3.25
enumerated conceptual domain
conceptual domain that is specified by a list of all its *value meanings* (3.48)

3.26
enumerated value domain
value domain (3.47) that is specified by a list of all its *permissible values* (3.38)

3.27
data element value
value out of a set of *permissible values* (3.38) pertaining to a *data element* (3.21)

3.28
definition
representation of a *concept* (3.15) by a descriptive statement which serves to differentiate it from related concepts

[SOURCE: ISO/IEC 11179-3:2013, 3.2.29]

3.29
identifier
<metadata registry> sequence of characters, capable of uniquely identifying that with which it is associated, within a specified *context* (3.19)

Note 1 to entry: A name should not be used as an identifier because it is not linguistically neutral.

3.30
metadata
data that defines and describes other *data* (3.20)

[SOURCE: ISO/IEC 11179-3:2013, 3.2.74]

3.31
metadata item
instance of a *metadata object* (3.32)

3.32
metadata object
object type defined by a metamodel

3.33
metadata registry
MDR
information system for registering *metadata* (3.30)

[SOURCE: ISO/IEC 11179-3:2013, 3.2.78]

3.34
name
primary means of identification of objects and *concepts* (3.15) for humans

3.35
object class
set of ideas, abstractions, or things in the real world that are identified with explicit boundaries and meaning and whose properties and behaviour follow the same rules

3.36
ontology
a conceptualisation of a domain

3.37
organization
<management> unique framework of authority within which a person or persons act, or are designated to act, towards some purpose

[SOURCE: ISO/TS 21089:2018, 3.97]

3.38
permissible value
expression of a *value meaning* (3.48) allowed in a specific *value domain* (3.47)

3.39
property
characteristic (3.12) common to all members of an *object class* (3.35)

3.40
registrar
representative of a *registration authority* (3.42)

3.41
registration
relationship (3.44) between an *administered item* (3.2) and the *registration authority* (3.42)

3.42
registration authority
RA
organization (3.37) responsible for maintaining a register

3.43
registration status
designation of the status in the *registration* (3.41) life-cycle of an *administered item* (3.2)

3.44
relationship
connection among model elements

3.45
unit of measure
<data> actual units in which the associated *values* (3.46) are measured

Note 1 to entry: The dimensionality of the associated conceptual domain shall be appropriate for the specified unit of measure.

3.46
value
data (3.20) *value* (3.46)

3.47
value domain
value set
set of *permissible values* (3.38)

Note 1 to entry: 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 to entry: The permissible values may either be enumerated or expressed via a description.

[SOURCE: ISO/IEC 11179-3:2013, 3.2.140 — modified, "value set" was added as a preferred term.]

Note 3 to entry: 'Value set' is Health Level 7 (HL7) terminology for value domain, as defined in Reference [13].

3.48
value meaning
meaning or semantic content of a *value* (3.46)

4 Framework for the management of metadata registry content

4.1 Overall approach

The ISO/IEC JTC 1/SC 32 International Standards portfolio on metadata provides a modular approach to the registration of metadata about a wide variety of models and the management of that registry. MDRMetamodel consists of two distinct components: common facilities defining metadata about content – metadata objects – with attributes for identification, naming, definition, classification, conceptualisation, administration and registration; and those defining content models for some of the metadata objects to be registered – concept systems, value lists and data elements. These common facilities are reused in MFIInformationModel for the registration of information models, MDRDatasets for datasets and MFIForms for form designs. A conceptual representation of the relationship between the components and standards is presented in Figure 1 where dependencies between standards are indicated by vertical layering. MDRMetamodel, MFIInformationModel, MDRDatasets and MFIForms cover a wide range of usecases, not all of which are necessary for application to healthcare.

The approach taken in this document is to distil minimal requirements from the portfolio that shall be met by compliant registries, present these as a consistent, unified metamodel, and then extend these minimal requirements for the specific requirements in healthcare. While the UML class diagram representation of the models is retained, a minimal subset of the UML is used in the presentation of the models, and the modelling style has been altered to facilitate translation into relational or document (XML/JSON) implementation. Nevertheless all models contained within this document are iso-semantic with those in the source standards, and registries compliant with this this document will additionally be able to claim basic compliance with MDRMetamodel, MFIInformationModel, MDRDatasets and MFIForms.

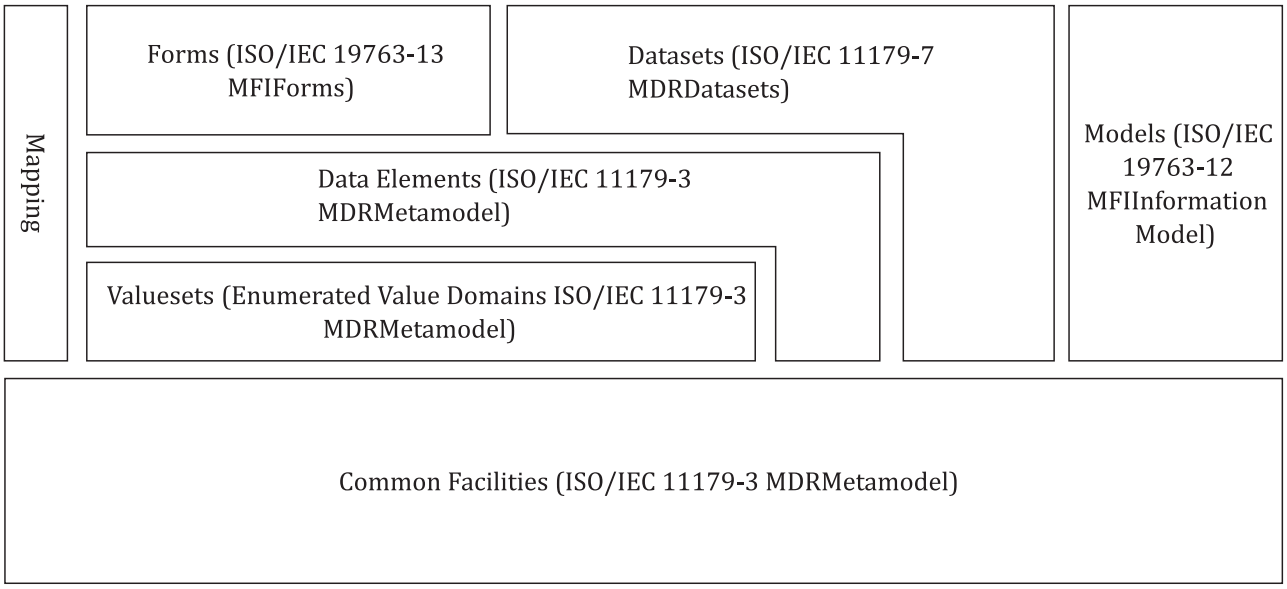

Figure 1 — Relationships between ISO/IEC JTC 1/SC 32 International Standards on metadata

The models in MDRMetamodel, MFIInformationModel, MDRDatasets and MFIForms are conceptual – there is no requirement that the specified model structures should be faithfully implemented in the actual schema of the registry/repository itself, nor any requirement to adopt either a relational, hierarchical or graph approach in their implementation, even though the literal interpretation of MDRMetamodel is most easily realised in a graph database. Relational implementations might implement separate schemas for administrative and semantic functions, provided the relevant administration record is linked to a class that is specified as a kind of administered item.

It is also worth noting that the basic data types described in MDRMetamodel, MFIInformationModel, MDRDatasets and MFIForms are those required for the implementation of the standard itself, not of the content of the registry/repository – MDRMetamodel has facilities for declaring specific datatypes of interest to a particular registry/repository's users, including those that are specific to healthcare information systems.

From this point on, registry and repository are used interchangeably unless specific reference is made. Implementers of simple repositories can choose not to implement those parts of the administrative elements of the common facilities which includes support the content creation lifecycle.

The Metamodel presented here is split into six sections: basic types; common facilities; data description; model description; schema registration; mapping, each of which are described below in textual form with illustrative, non-normative diagrams.

For a detailed discussion of how MDRMetamodel, MDRInformationModel have been restricted and extended in the derivation of this document, see Annex A.

4.2 Basic Types

4.2.1 General

Basic types contain foundational datatypes used in the metamodels described in this document.

4.2.2 BLOB (or binary large object)

A large block of binary data, typically an image or video file, that might have to be handled in a special way.

4.2.3 Boolean

A mathematical datatype associated with two-valued logic (see ISO/IEC 11404:2007, 8.1.1).

NOTE The notation and semantics for boolean are as described in ISO/IEC 11404.

4.2.4 Date

A datatype whose values are points in time to the resolution: year, month, and day (see ISO/IEC 11404:2007, 8.1.6).

4.2.5 Datetime

A datatype whose values are points in time to the resolution: year, month, day, hour, minute, second, and optionally fractions of seconds (see ISO/IEC 11404:2007, 8.1.6).

4.2.6 Integer

A mathematical datatype comprising the exact integral values (see ISO/IEC 11404:2007, 8.1.7).

NOTE Both the notation and semantics of the Integer datatype is as specified in ISO/IEC 11404:2007, 8.1.7.

4.2.7 Phone_Number

A phone number uniquely identifies a telephone line within a telephone network. The data structure of the Phone_Number data element shall conform to ITU-T E 164.

4.2.8 Postal_Address

A postal address enables the unambiguous determination of an actual or potential delivery point, usually combined with the specification of an addressee and/or a mailee.

4.2.9 String

A datatype comprising of a serial sequence of characters, bytes, integers, etc. See ISO 12639:2004, 4.1.11.

4.3 Common Facilities

4.3.1 General

Common facilities (see Figure 2) provides the model for an Administered Item, a registered item for which administrative information is recorded. Administered items are significant metadata entries that are individually identified, named, defined, classified, described by reference documents and have status with respect to registration and administration managed within the registry. Requirements for common facilities are derived from MFIMetamodel, but apply to all kinds of metadata item described in this document.