INTERNATIONAL STANDARD

ISO 17115

Second edition
2020-05

Health informatics — Representation of categorial structures of terminology (CatStructure)

Reference number
ISO 17115:2020(E)

© ISO 2020



COPYRIGHT PROTECTED DOCUMENT

© ISO 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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 ..... **iv**
Introduction ..... **v**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
3.1 General 1
3.2 Relations between concepts 3
4 Categorial structure description **4**
4.1 General 4
4.2 Precise goal of the categorial structure 4
4.3 Generic concepts 4
4.4 Type of characteristic 6
4.5 Generic concept relations 7
Bibliography **8**

© ISO 2020 – All rights reserved

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*.

This second edition cancels and replaces the first edition (ISO 17115:2007), which has been technically revised. The main changes compared to the previous edition are as follows:

— change in title and expansion of scope;

— reference to the ISO 17117 series which has further developed terms, definitions and characteristics of healthcare terminologies;

— inclusion of requirements for categorial structures to support the key purpose of standardization and clarification of terminologies in healthcare.

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.

© ISO 2020 – All rights reserved

Introduction

Health terminology is complex and multifaceted, more so than most language domains. It has been estimated that between 500 000 and 45 million different concepts are needed to adequately describe conditions of patients and populations, actions in healthcare and related concepts, such as biomedical molecules, genes, organisms, technical methods and social concepts.

To adequately represent and especially to process this complexity, simple coding schemes are inadequate and formal multidimensional concept representation systems are better suited. The differences in terminology are understandable as this kind of work is highly interdisciplinary and integrates knowledge from linguistics, philosophy, informatics and health sciences, and there is room for misunderstanding between disciplines.

Several such formal systems exist but systems and underlying philosophy are described in different ways. The system itself can, for example, be called an ontology, medical entity dictionary, coding and reference model or reference terminology.

Computer-based processing and interchange of medical or clinical information uses various kinds of terminological systems of concepts to represent that information, such as controlled vocabularies, classifications, nomenclatures, terminologies and thesauri, with or without coding schemes.

The specific terminological issues in the field of health informatics are the following:

a) a large number of different terminological systems are available in different clinical specialties;

b) a large overlap among the subject fields involved;

c) a large number of codes and rubrics, typically in the order of magnitude of 10 000 to 100 000 entries, in commonly used terminological systems;

d) an increasing need for re-use of coded data in different health-care contexts;

e) polysemy across and sometimes within clinical specialties.

The integration of computer-based medical records and administrative information systems design is supported by a uniform way to represent the meaning of medical concepts using terminologies. Such uniformity ensures that the EHR receiving a message will catch the meaning intended by the sender EHR and not just the string of characters used to represent that meaning. It is not possible to impose a rigidly uniform standardized natural language clinical terminology on healthcare providers due to the need for different levels of specificity based upon the clinical use case.

This document is intended to be read in conjunction with ISO 17117 and contains references to Reference [16].

A domain specific semantic model has been envisioned and applied in a series of specific European standards (EN) and International Standards (ISO) on various subject fields to describe a set of categorial structures. This is partially important where subject fields overlap. There are more than 10 different International Standards on categorial structures and this number is increasing every year.

Potential uses for this document are to

a) describe formal definitions, parts of definitions and how they are related, and

b) describe patterns for concept representation in a particular domain.

This document also

a) facilitates the construction of new terminological systems in a regular form which will increase their coherence and expressiveness,

b) facilitates maintenance of terminological systems,

© ISO 2020 – All rights reserved

c) increases consistency and coherence of existing terminological systems,

d) allows systematic cross-references between items of different types of terminological systems,

e) facilitates convergence among terminological systems and makes explicit the overlap between different health care domains terminological systems,

f) provide elements for negotiation about integration of different terminological systems into information systems between the respective developers, and

g) enable the systematic evaluation of terminological systems.

The target groups for this document are

a) developers of concept representation systems for different health care domains,

b) developers of standards for concept representation, especially those describing domain concept models,

c) information modellers, knowledge engineers, and standards developers building information models for health information systems such as electronic health records and decision support systems,

d) developers of information systems that require an explicit system of concepts for internal organization, data warehouse management and middleware services,

e) designers of specialized standard healthcare terminological categorial structures,

f) developers of healthcare terminological systems including classifications and coding systems,

g) producers of services for terminological systems and designers of software including applications for natural language processing,

h) information modelers, knowledge engineers, and standards developers building models for health information management systems,

i) developers of information systems that require an explicit system of concepts, and

j) developers of mark-up standards for representation of healthcare documents.

This document is informed by other standards, with clarifications and examples appropriate to healthcare.

This document has been developed for use as an integrated part of computer-based applications and for the electronic healthcare record. It would be of limited value for manual use.

© ISO 2020 – All rights reserved

Health informatics — Representation of categorial structures of terminology (CatStructure)

1 Scope

The purpose of this document is to

a) define a set of basic concepts required to describe formal concept representation systems, especially for health sciences,

b) describe representation of concepts and characteristics, for use especially in formal computer-based concept representation systems,

c) describe the characteristics which synthetically describe the organisation and content of a terminological system in health,

d) support the development of specific standards on categorial structures for particular healthcare subject fields with the minimum requirements to support meaningful exchange of information.

This document is not suitable for, or intended for use, by individual clinicians or hospital administrators. It is not the purpose of this document to standardize the end user classification or to conflict with the concept systems embedded in national practice and languages.

Topics considered outside the scope of this document include

— enumeration of axiomatic concepts and semantic links, and

— detailed content of health terminology systems (classifications, nomenclatures or reference terminology of health concepts).

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 General

3.1.1
categorial structure
reference terminology model
minimal set of *domain constraints* (3.1.5) for representing *concept systems* (3.1.2) in a *subject field* (3.1.7)

3.1.2
concept system
system of concepts
set of *concepts* (3.1.8) structured according to the relations among them

© ISO 2020 – All rights reserved

3.1.3
type of characteristic
characterizing category
category of *characteristics* (3.1.4) which serves as the criterion of subdivision when establishing *concept systems* (3.1.2)

EXAMPLE The type of characteristics 'colour' includes being red, blue, green, etc. The type of characteristics 'material' includes made of wood, metal, etc.

3.1.4
characteristic
abstraction of a property, of an object or of a set of objects

[SOURCE: ISO 1087:2019, 3.2.1, modified — added "of an object or of a set of objects".]

3.1.5
domain constraint
sanction rule prescribing the set of sanctioned *characteristics* (3.1.4) that are valid to specialize a *concept* (3.1.8) in a particular *subject field* (3.1.7)

EXAMPLE "Infection possibly hasLocation SkeletalStructure" describes that an infection in a particular context can be located in a structure that is a kind of skeletal structure.

Note 1 to entry: The rule describes the set of sanctioned characteristics by combining the *semantic link* (3.2.5) and the characterizing *generic concept* (3.1.6) it links to, possibly by enumeration of the concepts in the characterizing generic concept.

Note 2 to entry: Different levels of sanctioning are possible, e.g. conceivable, sensible, normal, usually InTheContextOf, necessary.

3.1.6
generic concept
concept (3.1.8) in a *generic relation* (3.2.1) that has the narrower intension

[SOURCE: ISO 1087:2019, 3.2.19]

3.1.7
subject field
domain
field of special knowledge

Note 1 to entry: The borderlines of a subject field are defined from a purpose-related point of view.

3.1.8
concept
unit of knowledge created by a unique combination of *characteristics* (3.1.4)

[SOURCE: ISO 1087:2019, 3.2.7]

Note 1 to entry: Informally, the term 'concept' is often used when the intended meaning is 'concept representation'. However, this leads to confusion when precise meanings are needed. Concepts arise out of human individual and social conceptualizations of the world around them. Concept representations are artefacts constructed of symbols and often manifest in computer programs. Because they are artefacts, it is possible to be precise about the functioning and capabilities of concept representations.

Note 2 to entry: Concept representations are not necessarily bound to particular languages. However, they are influenced by the social or cultural context of use often leading to different categorizations.

© ISO 2020 – All rights reserved

3.2 Relations between concepts

3.2.1
generic relation
genus-species relation
relation between two *concepts* (3.1.8) where the intension of one of the concepts includes that of the other concept and at least one additional delimiting *characteristic* (3.1.4)

Note 1 to entry: A generic relation exists between the *concepts* (3.1.8) 'word' and 'pronoun', 'vehicle' and 'car', 'person' and 'child'.

Note 2 to entry: This relation is equivalent to 'parent-child' or 'is-a' relation. The child concept has the same intension as the parent concept and at least one additional delimiting characteristic. A generic relationship is also represented in X 'is-a' Y.

Note 3 to entry: The use of the term 'relation' from an English perspective means that this is describing the concept (i.e. this concept is a generic relation of another concept), not the relationship. However, 'generic relation' here means the relationship itself, not a concept.

3.2.2
partitive relation
whole-part relation
relation between two *concepts* (3.1.8) where one of the concepts constitutes the whole and the other concept a part of that whole

Note 1 to entry: A partitive relation exists between the *concepts* (3.1.8) 'week' and 'day', 'molecule' and 'atom'.

Note 2 to entry: This relation is different from *generic relation* (3.2.1). For example, a day is part of a week, but is not a specialization of a week, i.e. it is not a type of week.

3.2.3
hierarchical relation
hierarchical concept relation
relation between two *concepts* (3.1.8) which can be either a *generic relation* (3.2.1) or a *partitive relation* (3.2.2)

[SOURCE: ISO 1087:2019, 3.2.12, modified — addition of "relation between two *concepts* (3.1.8) which can be either a".]

EXAMPLE Relations such as 'stomach' ⇔ 'organ' and 'stomach' ⇔ 'body' are hierarchical relations, but the former is a generic relation and the latter is a partitive relation.

Note 1 to entry: *Partitive relations* (3.2.2) have several sub-types, such as "constitutional part of", "regional part of", "grouped by", and so on, which are also hierarchical types of relations.

3.2.4
associative relation
associative concept relation
pragmatic relation
non-hierarchical concept relation

[SOURCE: ISO 1087:2019, 3.2.23]

EXAMPLE The relation between disease "X" and virus "Y" is not a hierarchical but an associative relation. The relation exists under a certain theme of interest (e.g. "a disease and its causative agent") and explicitly recognized by virtue of experience.

3.2.5
semantic link
formal representation of a directed *associative relation* (3.2.4) or *partitive relation* (3.2.2) between two *concepts* (3.1.8)

EXAMPLE hasCause (with inverse isCauseOf).

© ISO 2020 – All rights reserved